

EXTENSIONS OF A MAXIMUM ENTROPY ESTIMATED MARKOV DECISION  
PROCESS IN THE UNITED STATES AGRICULTURAL ECONOMY

---

A Dissertation

presented to

the Faculty of the Graduate School  
at the University of Missouri-Columbia

---

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

DUSTIN J. DONAHUE

Dr. Wyatt Thompson, Dissertation Supervisor

DECEMBER 2013

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled:

EXTENSIONS OF A MAXIMUM ENTROPY ESTIMATED MARKOV DECISION  
PROCESS IN THE UNITED STATES AGRICULTURAL ECONOMY

Presented by Dustin J. Donahue,

A candidate for the degree of doctor of philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

---

Professor Wyatt Thompson

---

Professor Pat Westhoff

---

Professor Willi Meyers

---

Professor Doug Miller

---

Professor Felix Fritschi

## DEDICATION

The author dedicates this work...

*To his parents* for their never-ending and unfailing support of his endeavors;

*To his brother* for showing him the power of a man committed to duty;

*To Terry, Steve, Russ, and the Brethren of Chattanooga Lodge #199, F&AM* for providing him with guidance and fraternal love above and beyond that which he can ever return;

*To the CoMo Derby Dames* for exposing him to a world that he never knew existed and for providing him a large extended family far from home;

*To Ziad* for setting him on this path and believing he could succeed;

*To Waffles* for being his long-time, long-distance confidant and sounding board;

*To his friends and relatives* for reminding him from whence he came.

## ACKNOWLEDGEMENTS

I wish to express my gratitude to my advisor, Dr. Wyatt Thompson. Long were the discussions, near arguments, over points of my research that needed to be changed. His patience with my stubbornness is a testament to his character.

I would also like to thank my committee members for their comments and constructive feedback that they provided. Their insights have helped to make me a better economist, looking at all angles of an issue.

I am grateful to the United States Department of Agriculture's National Institute of Food and Agriculture for awarding me the National Needs Fellowship which provided me with the ability to pursue my doctoral degree. I am likewise grateful to the University of Missouri Department of Agricultural and Applied Economics for its support.

~ Dustin J. Donahue

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	ii
LIST OF FIGURES .....	v
LIST OF TABLES .....	vii
ABSTRACT .....	viii
I. INCORPORATING DYNAMIC LAND USE INTO A PARTIAL EQUILIBRIUM MODEL: TEST CASES IN MISSOURI AND IOWA	
1. Introduction.....	1
2. Prior Literature.....	2
3. Markovian Land Use Decision Process .....	6
4. Partial Equilibrium Model .....	9
5. Data and Scope of Model.....	11
6. Additional Specifications.....	12
7. Model Constraints.....	16
8. Model Validation.....	17
9. Simulation Results: Constant Coefficient vs. MDP Comparison .....	19
10. Conclusion .....	22
11. References.....	24
12. Appendix: Figures and Tables.....	28
II. AN ECONOMIC EXAMINATION OF MIDWESTERN WARM SEASON GRASS AREA USING SATELLITE IMAGING DATA	
1. Introduction.....	34
2. Prior Literature.....	36
3. Markovian Land Use Decision Process .....	39
4. Data: WSG Returns.....	44
5. Data: Other Returns .....	45
6. Data: Acreage Shares .....	46
7. Data: Prior Probabilities.....	47
8. Spatial Correlation Issues .....	48
9. Constrained and Unconstrained Models .....	50
10. WSG Returns Shock and Area Response.....	53
11. Corn Net Returns Shocks and Area Response .....	56
12. Conclusion .....	58
13. References.....	61
14. Appendix: Figures and Tables.....	65

III.	A DYNAMIC ESTIMATION OF AN ILL-POSED US FEED DEMAND SYSTEM WITH DISTILLER'S DRIED GRAINS	
1.	Introduction.....	83
2.	Prior Literature.....	84
3.	Markovian Feed Use Decision Process.....	90
4.	Data: Feed and Residual Use.....	94
5.	Data: Explanatory Variables.....	96
6.	In Sample Dynamic Validation.....	97
7.	Model Elasticities.....	98
8.	Projection Data.....	103
9.	Projection Shocks and Results.....	104
10.	Conclusion.....	105
11.	References.....	107
12.	Appendix: Figures and Tables.....	111
VITA.....		119

## LIST OF FIGURES

Figure	Page
1.1	Illustration of the MDP-SPE Model Framework .....32
1.2	Markov Decision Process Model and Constant Coefficient Model Corn and Soybean Acres Planted Comparison to Historical Data, Iowa, 1996-2009 .....32
1.3	Markov Decision Process Model and Constant Coefficient Model Corn and Soybean Acres Planted Comparison to Historical Data, Missouri, 1996-2009 .....33
2.1	Midwestern Land Covers Observed by MODIS, Indication of Presence for Six Out of Ten Years, 2000-2009 .....65
2.2	Corn Net Returns to Corn Area Planted Elasticities, County Level Aggregation, Unconstrained .....66
2.3	Corn Net Returns to Corn Area Planted Elasticities, CRD Level Aggregation, Unconstrained .....66
2.4	Pastureland Rental Rate to Wheat Area Planted Elasticities, County Level Aggregation, Unconstrained .....67
2.5	Pastureland Rental Rate to Wheat Area Planted Elasticities, CRD Level Aggregation, Unconstrained .....67
2.6	Soy Net Returns to Corn Acres Planted Elasticities, CRD Level, Unconstrained .....68
2.7	Pastureland Rental Rate to Warm Season Grass Acres Elasticities, CRD Level, Unconstrained .....68
2.8	Net Returns Elasticities, CRD Level, Constrained Model.....69
2.9	Historical Warm Season Grass Area by CRD as a Percent of Total Land, 2000-2009 Average .....70
2.10	Change in WSG Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured in 1000 Acres.....71
2.10a	Change in WSG Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....71
2.11	Change in Wheat Area in Response to Doubled WSG Returns , 2001-2009 Average, Measured in 1000 Acres .....72
2.11a	Change in Wheat Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....72
2.12	Change in Soybean Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured in 1000 Acres .....73
2.12a	Change in Soybean Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....73
2.13	Change in Corn Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured in 1000 Acres.....74
2.13a	Change in Corn Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....74
2.14	Change in Corn Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres .....75

2.14a	Change in Corn Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....	75
2.15	Change in Warm Season Grass Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres .....	76
2.15a	Change in Warm Season Grass Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....	76
2.16	Change in Soy Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres.....	77
2.16a	Change in Soy Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....	77
2.17	Change in Wheat Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres .....	78
2.17a	Change in Wheat Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....	78
2.18	Change in Corn Area in Response to Higher Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres .....	79
2.18a	Change in Corn Area in Response to Higher Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres .....	79
2.19	Histograms of Changes in Area Allocated to Each Crop as a Percent of Total Area by Scenario, +/- 21.9% in Corn Net Returns, 2001-2009 Average.....	80
3.1	DDG Production 1992 – 2011 .....	111
3.2	Corn Feed and Residual Use as a Percent of Total Feed and Residual Use, Historical Amounts versus Dynamic Model Estimates, 2003-2012 .....	111
3.3	Soy Meal Feed and Residual Use as a Percent of Total Feed and Residual Use, Historical Amounts versus Dynamic Model Estimates, 2003-2012 .....	112
3.4	DDG Feed and Residual Use as a Percent of Total Feed and Residual Use, Historical Amounts versus Dynamic Model Estimates, 2003-2012 .....	112
3.5	DDG Price to Feed and Residual Use (as a Percent of Total Feed and Residual Use) Elasticities by Feedstock, 2003-2012 .....	113
3.6	DDG Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2013-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks .....	113
3.7	Corn Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2013-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks .....	114
3.8	Soy Meal Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2013-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks .....	114
3.8a	Soy Meal Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2010-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks .....	115
3.9	Soy Meal Feed and Residual Use (as a Percent of Total Feed and Residual Use) to DDG Price Elasticities, 2010-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks .....	115



## LIST OF TABLES

Table	Page
1.1	Markovian Land Transition Probabilities in Iowa, 1996-1997.....28
1.2	Demand Equation Price Elasticities.....28
1.3	Comparison of Net Returns to Acres Planted Elasticities with Measures of Fit (Mean Absolute Errors), Unconstrained MDP Models versus MDP Model Constrained with Coefficient Constraints and/or Non-Uniform Prior Probabilities .....28
1.4	Iowa Imposed Non-Uniform Prior Probabilities for Constrained Markovian Model.....28
1.5	Comparison of Historical Out of Sample Data with MDP-PE and CC-PE Model Base Scenario Estimates, 2006-2009 .....29
1.6	Differences between Estimated Markovian and Constant Coefficient Model Base Scenarios, 2006-2009, Constant Coefficient Model Shock Estimates Subtracted from Markov Model Shock Estimates .....30
1.7	Markov Model Scenario Differences, 3.0 Billion Gallon Shock Minus Baseline Scenario, 2006-2009 .....30
1.8	Constant Coefficient Model Scenario Differences, 3.0 Billion Gallon Shock Minus Baseline Scenario, 2006-2009 .....31
1.9	Difference of Differences Across the Markov and Constant Coefficient Models, Constant Coefficient Model Shock Difference Estimates Subtracted from Markov Model Shock Difference Estimates .....31
2.1	Example Markovian Land Transition Probabilities, 2001 – 2002.....81
2.2	Shares and Expected Net Returns by Crop, Averaged Across Selected CRDs and All Sample Years .....81
2.3	Prior Markov Probabilities Implemented in Model.....81
2.4	Percent Mean Absolute Error by Crop, Model Aggregation Comparison.....81
2.5	Model Percent Mean Absolute Errors by Crop, Model Constraint Comparison .....82
2.6	Regional Acreage in Million Acres by Crop Across Base and Shock Scenarios, 2001-2009 Average.....82
3.1	2003/04 – 2004/05 MDP Transition Probabilities .....116
3.2	Feedstock Sample Mean and Standard Deviation in Million Tons .....116
3.3	Explanatory Variables, Units, Means, and Standard Deviations .....116
3.4	Percent Mean Absolute Errors by Crop, 2003-2012.....116
3.5	T-Statistics of Model Coefficient Estimates by Feedstock.....117
3.6	Model Estimate Elasticities, 2003-2012 Average.....117
3.7	Feedstuff Substitution Elasticities.....117
3.8	FAPRI Baseline Elasticity-Analogous System Impact Multipliers for DDG Feed Use .....117
3.9	Feed Use Elasticities by Feedstock, 1992-2010 .....118

EXTENSIONS OF A MAXIMUM ENTROPY ESTIMATED MARKOV DECISION  
PROCESS IN THE UNITED STATES AGRICULTURAL ECONOMY

Dustin J. Donahue

Dr. Wyatt Thompson, Dissertation Supervisor

ABSTRACT

With an increase in the US focus on biorenewable energy, more forces are competing for agricultural land. This structural change in the agricultural economy warrants re-examination of the relationships between agricultural production decisions and the factors which influence those decisions. However, relevant data may be limited. To address this issue, a maximum entropy estimated Markov decision process model (MDP), a model ostensibly robust with limited data, is employed to examine agricultural decisions in three studies.

First, the question of the MDP's application to endogenous price changes is addressed by incorporating the MDP into a structural partial equilibrium model examining corn and soybean production in Iowa and Missouri from 1995-2005. This model is compared to a calibrated constant coefficient model and shocked to examine performance differences. The MDP was found to be more responsive to changes in price than a traditional model, although constraints on the model estimates were required to cause the model to follow economic response expectations.

Second, the MDP was applied to a newly acquired satellite imaging dataset showing warm season grass (WSG) area, a possible cellulosic ethanol feedstock, in the Midwestern US from 2001-2009, comparing the relationships between WSG, corn, soy, and wheat. The model proved problematic with large datasets, but showed the possibility

of WSG competing with traditional crops for area, responding to shocks in both its own price, and that of corn.

Third, the MDP was applied to US feed and residual usage. Because of the increase in ethanol production, dried distiller's grains (DDG) production has increased, creating a more available alternative livestock feed. DDGs, corn, and soy meal were examined from 2003-2012, and future use was projected. DDGs were found to compete with corn, but their soy meal relationship was unclear. DDG use is expected to level off and decrease slightly over the next ten years.

# **I. Incorporating Dynamic Land Use into a Partial Equilibrium Model: Test Cases in Missouri and Iowa**

## **I.1 Introduction**

In recent years, the United States has increased the focus it places on biofuel policy as a part of its overall energy policy. Specifically, the Energy Independence and Security Act (EISA) of 2007 includes mandates for the use of several categories of biofuel, some of which can be met using traditional (starch based) and cellulosic and agricultural waste based biofuels over the next decade. This mandate (which may or may not be waived at the discretion of the Environmental Protection Agency) could significantly change the manner in which U.S. land is allocated for crop production for the next decade. Even as a dedicated energy crop is developed to reduce price volatility transmitted from the oil market to traditional crop prices, the new crop will compete with traditional crops for land. Even setting aside the potential for new crops, the impact of biofuel on the mix of crops grown in the U.S. has been a matter of debate, with increased pressure in corn markets apparent in area shifts in recent years. Because of this impact, there is benefit in knowing how different factors will affect land allocation: understanding these interactions will allow policy makers to make better decisions.

Many prior studies exist on land use allocations (LUA). Often, LUA models incorporate a set of constant parameters which, with other data, allow for the estimation of land use. For instance, the Food and Agricultural Policy Research Institute (FAPRI) utilizes a structural partial equilibrium (PE) model to project U.S. commodity and crop production. A PE model is built with a series of linear or non-linear equations modeling both supply and demand for various commodities at an aggregated level. However,

constant coefficient (CC) models necessarily assume that relationships between data and land use remain at a single level throughout time. The question arises whether or not constant parameters are appropriate for LUA models, as opposed to a dynamic coefficient model, in which the relationship between factors determining land allocation and the allocation decision change over time. One method to introduce dynamic coefficients is to assume that land use follows a first order Markov decision process (MDP). It is a dynamic process in which the current land allocation is a function of the last period's allocation and a set of non-stationary transition probabilities. These probabilities are estimated based on data which affect the decision process. This allows a greater freedom of movement in a model, which may allow it to reflect reality more accurately.

However, traditional PE models have the ability to determine price endogenously and solve recursively for changing market conditions. Prior Markov models have not incorporated such price endogeneity. Combining these two capabilities may provide a more robust model, thereby serving as a better tool for analysis. The questions this research intends to answer are 1) Can a land use model incorporating an MDP endogenously determine price, and 2) If so, how does it compare with a traditional purely constant coefficient model? To answer these questions, this research incorporates a MDP into a PE model.

## **I.2 Prior Literature**

One of the possible extensions of this research is examining the impacts of biofuel policy on land use. With the increasing importance of biofuels in the American agricultural economy, literature concerning such impacts on land use and other areas has increased in frequency. Devadoss and Bayham (2010) examine the effects of reducing

crop subsidies on biofuel production. Dicks et al. (2009) examine the amounts of various feedstocks necessary to satisfy the EISA mandate. Miller and Coble (2011) study different possible policy options available to lawmakers and the impact of those proposals on the Southeastern U.S. Kim, Schaible, and Daberkow (2010) focus on the impacts of U.S. biofuel policy on international energy markets. While this list of biofuels related economic literature is far from exhaustive, it makes clear that there are a great many questions to answer about the changing bioenergy economy. This research focuses on developing a new model to better address some of them.

Structural models have been used to solve for land use and agricultural production for several decades. An early use of the structural model comes from Womack (1976). He models U.S. agricultural production at a national level, incorporating both crop and cattle production and consumption. Shumway (1983) studies Texas field crop production, but his analysis focuses mainly on the supply side of the equation and the cost of inputs. Adams (1994) follows Womack, and includes a regional breakdown of the national model. A prominent current use of this type of model can be found in the FAPRI Baseline, as mentioned earlier (Westhoff and Brown 2009).

The earliest use of a MDP to estimate changes in land use is found in Burnham (1973). The author assumes that land use changes in the Southern Mississippi Alluvial Valley can be estimated by a stationary first order Markov chain. However, he states in the footnotes that assuming stationary transition probabilities may be too restrictive for a land use change model. Burnham's conclusion is supported by Hallberg (1969), who studies frozen dairy products in Pennsylvania using a non-stationary MDP. Hallberg uses multiple regression techniques to test the hypothesis that non-stationary Markov

transition probabilities have better predictive capability than stationary probabilities. MDPs are used in far more than land use, however. Again, Hallberg's study is on firm size in the dairy industry, not land use. Other studies using MDPs for econometric analysis include Adelman's analysis of the distribution of firm sizes in the iron and steel industry (1958); Paap and Van Dijk's analysis of income and consumption in the U.S. (2003); Kelley and Weiss' study of population migration based on wage differentiation (1969); Miller and Plantinga's paper analyzing land use changes in Iowa (1999); and Lubowski, Plantinga, and Stavins' use of a nested logit model to analyze national non-federal land use (2008).

As stated earlier, other models are utilized in economics to determine land use allocation. One such method is an iterative linear program (LP). In an LP, a linear production function is optimized subject to constraints such as labor costs, crop net returns, and total acreage constraints. Heady (1954) provides a discussion of the logic and advantages of using linear programs in agricultural econometrics. Tompkin (1958) uses linear programming to determine the optimal combination of production activities on a livestock farm. A different type of LP model is used by De La Torre Ugarte et al. (2003) and Ray et al. (1998) in the implementation of the POLYSYS agricultural policy model. However, LP models may result in corner solutions, so transition constraints may be necessary.

Another model used to estimate land use shares and transition probabilities is the multinomial logit model (MNL) (Theil 1969). The MNL is used by McRae (1977), among others. The probabilities associated with a change in state are estimated using a logistical form. This functions similarly to a Markov chain, but disregards the prior state,

estimating changes in land use solely as a function of exogenous data. Such use can be seen in Wu and Segerson (1995); Hardie and Parks (1997); and Ahn, Plantinga, and Alig (2000), for instance. In addition, Lubowski, Plantinga, and Stavins (2003, 2008) note that the Independence from Irrelevant Alternatives (IIA) property of MNLs may preclude otherwise optimal choice behaviors. They, in addition to Lubowski (2002), use a model known as a nested logit model (NLM) to address this shortcoming of the MNL. The NLM separates decision states into subgroups or “nests” of similar qualities, differentiating them based on degree of substitutability. The nests Lubowski, Plantinga, and Stavins use include urban, non-farm (comprised of forest and range land), and farm (comprised of cropland, Conservation Reserve Program land, and pasture land). The advantage of the NLM is that it imposes IIA within nests, but not across nests, relaxing the choice restrictions. However, because the nests of a NLM are based on substitutability, it may be less efficient at explaining land use change between crops. The mixed logit model (MLM) is another method of relaxing the IIA restriction. The MLM is given by McFadden and Train (2000), and takes the choice specific variables of conditional logit and the choice-independent variables from the MNL to create a mixed model in which additional choices change the relative probabilities of the existing states.

However, the MLM still requires a problem that is well posed: i.e. the observations available exceed the number of unknowns. In the case of the introduction of a new crop for cellulosic biofuel feedstocks, the known data are very limited, and the problem may be ill-posed. Golan, Judge, and Miller (1996) discuss the use of general maximum entropy to address the issue of an ill-posed problem. From this discussion and other existing land use literature, a model that estimates a set of coefficients directly



linking explanatory state variables (both state dependent and independent to avoid IIA) with decision variables, as well as incorporating data from the state of the land in the prior period may be derived to provide a better fit. It is also desirable to utilize an econometric model that minimizes choice restrictions on the part of the decision maker. One such model that fits those criteria is the MDP model.

### **I.3 Markovian Land Use Decision Process**

This research examines the dynamics of land use changes by explaining the planting decisions of farmers in the area of study. Following Ahn, Plantinga, and Alig (2000), each farmer in region  $i$  ( $i = 1, \dots, I$ ) is assumed to plant a sequence of crops on land that he manages that maximizes the present discounted value of expected net returns

$$(1.1) \quad \max_{\delta_{j,k,t}} \sum_{t=0}^{\infty} \tau^t \cdot E[NR(\delta_{j,k,t}, \mathbf{X}_{h,i,t}) + \varepsilon_t(\delta_{j,k,t})]$$

where  $\tau$  is a constant discount factor,  $\delta_{j,k,t}$  represents a decision to allocate land currently used for crop (discrete state variable)  $j$  in time  $t-1$  to crop  $k$  in time  $t$ ,  $\mathbf{X}_{h,i,t}$  represents a  $T \times H$  matrix of observable data, including expected crop net returns and a trend variable, and  $\varepsilon_t$  represents unobserved variables. Beyond the limit of a farmer making decisions only for the land he owns or controls, this maximization is unconstrained. Because of the relative difficulty involved with obtaining individual farmer planting decisions, a model which utilizes aggregated data is desirable. Therefore, in this paper, acres planted aggregated at the state level are used to represent the sum of all planting decisions (Golan, Judge, and Miller 1996). However, this research recognizes that there exist idiosyncratic or agent-specific components of a farmer's planting decision, which are unobservable to the econometrician. These components are represented by the

unobservable state variable  $\varepsilon$ . Over time, a farmer's planting decisions are assumed to follow a first order MDP. The acreage planted to a given crop is a function of the crop planted in the prior period and a non-stationary transition probability, noted  $\pi_{i,j,k,t}$ . That is, there is a non-observable  $J \times K$  (where  $J=K$ ) matrix of probabilities for each state and time period, transitioning from crop  $j$  to crop  $k$  from period  $t-1$  to period  $t$ .

To illustrate, Table 1.1 shows the probabilities for land use transition in Iowa from 1996-1997 estimated from the model explained below. If an acre of land was planted with corn in 1996, it has approximately a 57% chance of staying in corn in 1997, a 37% chance of being planted with soybeans, and a less than 6% chance of being used in another category. Here, these numbers provide a visual example of the nature of Markovian transition probabilities; as results, they will be discussed later. Markovian transition probabilities row-sum to one and are hypothesized to be affected by explanatory variables including crop prices. Through this interaction, the variables are explanatory of changes in crop land planting patterns.

The model is derived using maximum entropy (ME), following Golan and Vogel (2000); Golan, Judge, and Miller (1996); and Miller and Plantinga (1999). The ME method for estimating the model of Markov transition probabilities is set forth by Jaynes (1957). The objective of the ME method is to select the probabilities that use the least information (fewest assumptions) to estimate the probabilities while still satisfying the constraints. This, in turn, assumes the farmer has the greatest amount of choice possible. Shannon's (1948) entropy measure is used to measure the amount of information needed to estimate the coefficients (Miller and Plantinga 1999, Golan and Vogel 2000). The primal objective function determining the optimal transition probabilities is

$$(1.2) \quad \begin{aligned} \max_{\pi_{i,j,k,t}} S_i &= \sum_{t=1}^T \sum_{j=1}^J \sum_{k=1}^K \pi_{i,j,k,t} \cdot \ln(\pi_{i,j,k,t}) \\ \text{s.t.} \quad \sum_{t=1}^T \mathbf{X}'_{i,t} \left( y_{i,k,t} - \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t} \right) &= \bar{\mathbf{0}} \end{aligned}$$

where  $\bar{\mathbf{0}}$  is an  $H$  size vector of zeroes and  $y_{i,k,t}$  is the share of acreage planted to crop  $k$  in area  $i$  during time  $t$ . When applied to the estimating equations (the set of constraints), the solution to the problem takes the form

$$(1.3) \quad \hat{\pi}_{i,j,k,t} = \frac{q_{i,j,k,t} \cdot \exp\left(y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \hat{\lambda}_{h,i,k}\right)}{\sum_{k=1}^K q_{i,j,k,t} \cdot \exp\left(y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \hat{\lambda}_{h,i,k}\right)},$$

where  $\hat{\lambda}_{h,i,k}$  is the optimal Lagrangian multiplier associated with explanatory variable  $h$  and crop  $k$ , and  $q_{i,j,k,t}$  is a conditional probability, which may be adjusted by the analyst to represent information from before the planting decision that may bias the farmer's decision. By focusing on the dual of the primal problem, an unconstrained equation to determine the optimal multipliers takes the form of

$$(1.4) \quad \begin{aligned} \max_{\lambda} M_i(\bar{\lambda}) &= \sum_{t=1}^T \sum_{h=1}^H \sum_{k=1}^K y_{i,k,t} x_{h,i,t} \lambda_{h,i,k} \\ &- \sum_{t=1}^T \sum_{j=1}^J \ln \left[ \sum_{k=1}^K q_{i,j,k,t} \exp \left( y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \lambda_{h,i,k} \right) \right] \end{aligned}$$

One  $k$  category is kept as a residual, with its multipliers assumed to be zero. This implies that the residual land is unaffected by the explanatory data. Because of the additive nature of the land categories, the residual solution is implicit when all other categories are estimated. After the multipliers are estimated, they are used with the other variables to determine the transition probabilities in (1.3). The transition probabilities are then applied to the prior year's crop plantings to determine current plantings:

$$(1.5) \quad \hat{y}_{i,k,t} = \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t}.$$

With the acreage shares estimated, elasticities can be calculated to show the effect of a one percent change in the explanatory variables on the transition probabilities. (Miller and Plantinga 1999)

$$(1.6) \quad \omega_{h,i,j,k,t} = x_{h,i,t} \cdot y_{i,j,t-1} \left[ \hat{\lambda}_{h,i,k} - \sum_{k=1}^K \hat{\lambda}_{h,i,k} \cdot \pi_{i,j,k,t} \right].$$

Using this elasticity, one can derive the acreage transition elasticities, measuring the change in the acreage allocation from a one percent change in the explanatory variables, as

$$(1.7) \quad \kappa_{h,i,k,t} = \hat{y}_{i,k,t}^{-1} \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t} \cdot \omega_{h,i,j,k,t}.$$

Using the appendix in Miller and Plantinga (1999) as a guide, the covariance matrix of the coefficients ( $\hat{\lambda}_{h,i,k}$ ) conditional on the explanatory variables can be estimated by

$$(1.8) \quad \left( (\mathbf{X} \otimes \mathbf{I}_{(K-1)^2})' \boldsymbol{\Sigma}_i (\mathbf{X} \otimes \mathbf{I}_{(K-1)^2}) \right)^{-1}$$

such that  $\mathbf{I}$  is the identity matrix, and  $\boldsymbol{\Sigma}_i$  is a  $T(K-1)^2 \times T(K-1)^2$  matrix where the diagonal elements are  $\pi_{i,j,k,t}(1 - \pi_{i,j,k,t})$  and the off diagonal elements are  $\pi_{i,j,k,t}(-\pi_{i,j,k,t})$ .

With the MDP defined, the focus shifts to defining the PE model which will endogenously determine price.

#### **I.4 Partial Equilibrium Model**

The PE model used here is a set of linear demand and price equations designed to take the land allocation decision from the MDP and allow it to interact with the market, endogenously determining price, and allowing that price to feed back recursively into the

variables affecting the farmer's planting decision. The following equations specify the relationships between simulated variables and the MDP.

$$(1.9) \quad \text{PlantedAcres}_{i,k,t} = y_{i,k,t} \cdot \text{TotalLand}_{i,t}$$

To reach actual acres planted to each crop, the shares are multiplied by the exogenous total amount of land available. Shares are used in estimation because of the MDP's probabilistic nature, preventing projections from being tied to the last year of estimation.

$$(1.10) \quad \text{HarvestAcres}_{i,k,t} = \text{PlantedAcres}_{i,k,t} \cdot \text{HHR}_{i,k,t}$$

Acres harvested for each crop in a given year are calculated by multiplying the acres planted by the average historical harvest rate for that crop. This accounts for area harvested for silage as opposed to grain, which is studied here, and may include factors like crop failure or area not harvested for other uses.

$$(1.11) \quad \text{Production}_{i,k,t} = \text{ActualYield}_{i,k,t} \cdot \text{HarvestAcres}_{i,k,t}$$

Actual yield is exogenous, so production is the product of the area harvested and the realized yield.

$$(1.12) \quad \text{Supply}_{k,t} = \sum_{i=1}^l \text{Production}_{i,k,t}$$

Total supply for the region is the aggregated production amounts of each state.

$$(1.13) \quad \text{Demand}_{k,t} = \beta_{01} + \beta_{11} \text{Income}_t + \beta_{21} \text{Price}_{k,t} + \beta_{31} \text{Price}_{l,t}$$

Demand is a function of national income and prices. Here,  $l$  indicates all discrete states (crops) in  $t$  such that  $k \neq l$ . Thus, both own and cross prices are included.

$$(1.14) \quad \text{Supply}_{k,t} = \text{Demand}_{k,t}$$

This equation clears the market and varies price to satisfy the constraint.

$$(1.15) \quad \text{Price}_{k,t} = f(\text{Price}_{i,k,t})$$

The market-clearing regional prices are linked back to the state-specific prices. By inserting the MDP supply response into the above equations, the model becomes a MDP-PE. The model for this paper is further defined by the scope and data.

### **I.5 Data and Scope of Model**

The objective of this research is to compare a MDP-PE model with a CC-PE model. Figure 1.1 provides an illustration of the overall framework of the model and the relationships between the variables. Area for two crops that are already produced commercially is estimated over a historical period and combined with a structural partial equilibrium model; then, an out of sample simulation is run with and without an external shock to compare the two models. Here, corn and soybeans are estimated in Iowa and Missouri for the years 1996-2005, with an out of sample simulation in 2006-2009. The residual category consists of all other land available in both states after corn and soybean acres are removed. While the residual category is included for estimation purposes, it is not a focus of this research: the residual category has no returns calculated for it. As such, one limitation of this study is that the focus on corn and soybeans is narrowly defined, with land associated with all other activities left as an “other”. However, as noted below, the treatment is consistent among the land use equations that are compared in the models presented here.

With this framework in mind, the area variable for the MDP-PE  $i = [\text{Iowa, Missouri}]$ , the crop category variables  $j,k = [\text{corn, soybeans, residual}]$ , time variable  $t = [1996, 1997, \dots, 2005]$  and the explanatory variable  $h = [\text{corn net returns, soybean net returns, intercept}]$ . State level data for corn and soybeans on acres planted, acres harvested, yields, fertilizer application rates, and farmgate prices for each marketing year

are taken from the National Agricultural Statistics Service (NASS). Total state land area is taken from the 2010 U.S. Census, held constant over time, and used to determine shares. Fertilizer price data for anhydrous ammonia (N), di-ammonium phosphate (P), and muriate of potash (K) are obtained at the regional level. The NASS regions used are Heartland and North Central. National corn and soybean farmgate prices, as well as soybean meal prices, are obtained from NASS as well. National personal income is obtained from the Economic Research Service (ERS). All prices and income are in real terms. Consumer prices and income are adjusted using the consumer price index from the Bureau of Labor Statistics, with farmgate and policy prices adjusted using the producer price index from the same.

## I.6 Additional Specifications

To simulate the model, a number of additional equations are necessary apart from the land allocation problem that was explained before.

$$(1.16) \quad Price_{Missouri,k,t} = \beta_{02} + \beta_{12} Price_{Iowa,k,t}$$

Prices for both corn and soybeans in Missouri were found to follow Iowa (the larger producer) almost perfectly. Therefore, the model solves for Iowa prices, and Missouri prices are taken as a function of them.

$$(1.17) \quad E[NR_{i,k,t}] = \max(FuturesPrice_{i,k,t} - Basis_{i,k}, PolicyFloorPrice_{i,k,t}) \\ \cdot E[Yield_{i,k,t}] - E[VariableCost_{i,k,t}]$$

Expected net returns are equal to the maximum of state-level futures price and a policy-set floor price times expected yield minus expected variable costs. The national futures price is gathered from the Chicago Mercantile Exchange, and, following crop insurance (Hofstrand and Edwards 2003) the average of daily settlement prices in February were

used to calculate the futures price used. Corn and wheat prices were based on December contracts, while soybean futures prices were based on November contracts. The futures price is adjusted by subtracting the average difference between the futures price of the chosen month and the state-level average farm price. Although this gap varies from year to year in ways that are assumed to be predictable, farmers are expected to incorporate the average difference between the price on the CME and the price they receive that is sustained over this period. The floor price is derived from the United States marketing loan program, but is assumed to run higher than the loan rate set by policy based on the observation that the marketing loan benefits plus market price typically exceeds the loan rate in those years where payments were made during the sample period.. The policy floor price was estimated as the loan rate plus the average percentage by which loan deficiency benefits plus state level market prices exceeded loan rates over the years where there were benefits during the estimation period of this study. Expected yield is a trend yield.

$$(1.18) \text{ FuturesPrice}_{i,k,t} = \text{FuturesPrice}_{k,t} - \text{StateFuturesDiff}_{i,k,t}$$

The national futures price is translated to the state level by subtracting the average difference between the current state marketing year price and the national futures price. Dickey-Fuller tests were conducted on the differences to determine stationarity. No unit root was found at the 1% level of significance for corn or soybeans in Iowa. Because Iowa prices drive the model, Missouri was not tested.

$$(1.19) \text{ FuturesPrice}_{k,t} = f(\text{Price}_{\text{Iowa},k,t-1})$$



Out of sample, the futures price is estimated as a function of the lagged marketing year price in Iowa. Iowa prices drive the model, and this equation closes the price expectation system.

$$(1.20) E[VariableCost_{i,(Corn,Soy),t}] = \sum_{Fertilizers=[N,P,K]} FertAppRate_{i,(Corn,Soy),t,(N,P,K)} * FertPrice_{(N,P,K),t}$$

Aggregate per-acre fertilizer costs are included as the variable costs for each crop. Other variable costs, including fuels, chemicals, and seeds, are not included for two reasons. First, fertilizer is assumed to explain an important part of the differences between corn and soybean costs. Second, fertilizer relates the key research question of this study because of how fertilizer requirements and crop rotation interact. For soybeans, the fertilizer costs are the per-acre application rates times the price.

$$(1.21) FertAppRate_{i,Corn,t,N} = \max\left(\frac{0.5, y_{i,Corn,t} - y_{i,Soy,t-1}}{y_{i,Corn,t}}\right) \cdot C2CAppRate + \min\left(\frac{0.5, y_{i,Soy,t-1}}{y_{i,Corn,t}}\right) \cdot S2CAppRate$$

For corn, the issue of soybean rotation arises. Farmers, to fix nitrogen in the soil, will plant soybeans after corn, thereby reducing fertilizer costs. This functional form assumes a 50% rotation of all available acres in state-level data, unless there is a portion of corn acreage that cannot be rotated. Given the recommended corn-over-corn and corn-over-soy application rates come from the Iowa State University Extension Service (Blackmer, Voss, and Mallarino 1997), correspondingly higher fertilizer costs are built in if state-level data suggest that some corn must be planted on area that was allocated to corn in the previous year. Current fertilizer prices and application rates are used because the farmer has knowledge of the prices he faces when he purchases fertilizer, as opposed to facing an expectation of price. It should be noted here that the issue of soil quality or

productivity is recognized as affecting the farmer's decision process (Orazem and Miranowski 1994, Thomas 2002, Miller and Plantinga 1999). However, because this model is aggregated at the state level, land quality differences are difficult to capture, and are assumed to reside in the idiosyncratic portion of the optimization problem.

$$(1.22) \quad \ln(Demand_{Corn,t}) = a0_{corn} + a1_{corn} * \ln(USCornPrice_t) \\ + a2_{corn} * \ln(USSoyPrice_t) + a3_{corn} * Trend_t$$

$$(1.23) \quad \ln(Demand_{Soy,t}) = a0_{soy} + a1_{soy} * \ln(USCornPrice_t) \\ + a2_{soy} * \ln(USSoyPrice_t) + a3_{soy} * Trend_t$$

Demand for corn is a function of its own farmgate price and the farmgate soy price, as well as intercepts and trends. Likewise, demand for soy is a function of its own farmgate price, the corn farmgate price and other terms. For demand, regional production is used as the dependent variable, but demand for crops grown in these two states implicitly includes all uses and all competing producers, domestic and foreign. Regional demand elasticities are based on the method applied for a model that included corn and soybean production in Illinois (Khanna et al. 2009), using with the same estimates of demand and supply, namely demand elasticities of -0.16 for corn and -0.59 for soybeans and supply elasticities of 0.2 for corn and 0.45 for soybeans, but taking into account the share of crop production in this two states (USDA/NASS 2012). Since (1.22) and (1.23) use a double log form, the demand in shown in (1.13) would be the exponential of (1.22) or (1.23).

Estimation comparison is restricted to validating the non-constant representation of supply, whereas the standard representations of demand and price-clearing equations are added to simulate the impact of an external shock in order to show the potential for differences in results if non-constant supply parameters are used in policy analysis. For

out-of-sample simulations, the model is calibrated to the final year of historical estimation. The elasticities are listed in Table 1.2. Own-price elasticities are negative and very elastic, as is judged appropriate given that the demands are for crops produced in only two states, and cross-price elasticities are positive and inelastic.

### **I.7 Model Constraints**

As written, the model should run unconstrained to choose the coefficients which best fit the data. Indeed, an unconstrained model is desirable because it may be considered more statistically valid, as it more closely reflects the relationships present in the data. However, in applied econometric work, constraints are often imposed if the results do not fit economic theory or if the data may not perfectly reflect the situation faced by the decision maker, as in the case of proxies. To see if such constraints were necessary, the model is run over the in-sample period with actual lagged acreage and explanatory data. The acreage elasticities (1.7) of the net returns variables were examined, as well as the structure of the Markov probability matrices (1.3). The initial run of the model produced elasticities which do not follow expected economic theory: own net return elasticities should be positive, and cross net return elasticities should be negative. Of the eight elasticities calculated, only the effect of corn net returns on soy acres had the expected sign.

To address this issue, two sets of constraints were imposed: the first was imposed on the value of the coefficients calculated in (1.4), the second on the prior probabilities  $q$  in (1.3). The elasticities were indirectly constrained by way of the coefficients themselves because elasticities are not explicitly represented in this estimation method, but are calculated ex post instead. These constraints were designed to direct the model as little as

possible while still ensuring the correct signs and relative magnitudes on own net returns elasticities. In Iowa, restricting soy net returns to corn acres as negative produced results which fit theory. In Missouri, corn net returns to corn area was constrained to be positive. These constraints are imposed to cause the model to follow the known competitive nature of corn and soybeans. The elasticities from both models, averaged over time, are shown in Table 1.3.

The second set of constraints were imposed on the probability matrix (1.3) by introducing a non-uniform set of prior probabilities ( $q$ ). While specific crop rotation data are difficult to obtain, Wallender (2013) shows that while some area stays in corn continuously, most is rotated to soybeans at least once in three years. Following this, the prior probabilities assumed a corn-corn-soy rotation of aggregate shares, with a very small possibility of land going into residual. Corn over corn probabilities were increased slightly to account for the small percent of land that does not rotate. These priors are presented in Table 1.4. Missouri prior probabilities were more difficult to determine because of the much broader array of crops grown there. Also, Missouri estimated probabilities of residual land going into corn and soy were very low without imposing priors, therefore Missouri was left with uniform priors.

### **1.8 Model Validation**

It should be noted that forcing the model to reflect theory did not come without cost. While the economics are more closely followed, the constraints impact the statistical validity of the errors of the coefficients derived from the covariance matrix in (1.8). For this reason, confidence intervals and t-tests are not provided for the coefficients. The question also arises as to how the constraints affect the ability of the model to estimate

land use changes, essentially a question of goodness of fit. Due to the dynamic nature of the MDP, traditional measures of fit (i.e.  $R^2$ ) are not statistically valid. But, a different measure of fit for the MDP may be presented:

$$(1.24) \quad MAE_{i,k} = \frac{1}{T} \sum_{t=1}^T \frac{abs(\hat{y}_{i,t,k} - y_{i,t,k})}{y_{i,t,k}}$$

MAE is the percent mean absolute error: the difference of the estimated acreage and the actual acreage for each crop, normalized by the actual acreage planted to that crop, averaged over time. As it measures deviation from historical data, a lower number indicates a better fit. Because  $y$  represents a share, as opposed to actual acreage, this estimate is a unitless percentage. Again, this does not have the same statistical application as a normal  $R^2$ , but it does provide an idea of how the model performs. The MDP in-sample MAEs are listed in Table 1.3. The constraints on the model decrease its ability to estimate land use changes with regards to corn acres, as shown by the higher MAEs in the constrained model. However, soy acres perform the same or better in both Missouri and Iowa.

The constraints have some impact on the model, but it may be easier to graphically validate the model and to see how it compares against historical acres planted to corn and soy, shown in Figures 1.2 and 1.3. The vertical double-dash line indicates the change from in-sample to out-of-sample estimation. Both follow fairly closely over time, but the model misses the changes in 2007 when corn and soy prices spiked. The graphs seem to suggest that the model catches the turns, but misses the magnitudes of the changes. Because it catches the turns in production, it may be useful to see how the model responds to a shock in prices. Therefore, the focus shifts to comparing the performance of a CC version of the model to the combined, price endogenous MDP-PE.

## **I.9 Simulation Results: Constant Coefficient vs. MDP Comparison**

To compare the MDP model to a CC model, a CC-PE model is created and calibrated in part to the constrained MDP-PE estimation results to ensure similar orders of magnitude. The CC-PE acreage response equals the elasticities calculated from the MDP-PE, averaged over the in-sample period. Intuitively, this approach suggests that an estimated CC-PE model over the same period might end up with the constant supply elasticities equal to the average values of the non-constant elasticities reported above. This calibration is chosen to keep the magnitude of response from the CC-PE model in line with that of the MDP-PE model. To provide a baseline for the differences in the two models, Table 1.4 shows the out-of-sample relevant data for each model and the actual observations. In comparing the averages of the base scenarios, the initial differences seem to follow theory: both the CC-PE and MDP-PE models allocate fewer acres to these crops when their own prices are lower, with the lower prices in both cases possibly a result of simulated demand weaker than historical demand during this period. However, the CC-PE allocates fewer acres to crops than the MDP-PE and has slightly higher prices. The acreage differences are relatively small, but this does provide an indication that, despite calibration, the models will respond differently. Table 1.6 shows these differences more explicitly. The difference between the two begins small in acres planted, but spreads apart as the simulation progresses in most categories. Another measure of performance of a structural model can be generated by simulating the results of an external shock.

To test the model further, a shock was applied in the first year of the simulation period. In keeping with a stylized applicability to biofuels, an exogenous increase in regional ethanol demand of 3.0 billion gallons in 2006, with an additional 500 million

gallons each year, was applied. A conversion rate of 2.7 gal/bu is assumed, based on the lower end of the range tested by Donner and Kucharik (2008). Distiller's grains are included as a percent reduction in total corn demand from the shock based on per-gallon distiller's grain production numbers from USDA (Hartman 2013), under the assumption that increased distiller's grains will compete with corn for feed demand, attenuating the increased demand for corn. Table 1.5 shows the differences between the two models across the two scenarios (the difference of differences). The shock is assumed to be unanticipated, so expectations of future crop prices are not affected by any probability that farmers might otherwise place on the price impacts of steadily increasing ethanol production from corn. Before the models are contrasted, one should examine the signs and magnitudes of the individual scenario differences to evaluate performance and ensure that the simulations follow economic theory.

Tables 1.7 and 1.8 show the shock scenario differences, subtracting the base scenario from the shock scenario. In 2006, the supply side parameters (acres planted and production) do not respond due to lagged response variables. However, the increase in ethanol demand causes an increase in the price of corn as the new demand is filled at the expense of other demands. The price of soybeans also increases as livestock owners and consumers shift consumption to soybean products as the price of corn rises. The CC-PE model has a higher increase in prices than the MDP-PE from the identical shock in quantity demanded, suggesting that the MDP-PE has a greater price elasticity of supply, which is to say an identical change in price across both models would evoke a stronger supply response in the MDP-PE. In 2007, the entire model picks up the shocks as the lagged increases in price cause increases in corn acreage planted to help fill the

exogenous demand shock. There are fewer acres planted to soy in Iowa in 2007, despite a higher soy price. This suggests a “bleedover” effect from the demand shock of the previous year. Acreage is taken out of soybeans, despite the fact that soybean prices have increased. This can be explained by the fact that the corn price has risen higher than the increase in the price of beans in percent terms; so acreage is taken away from soybeans.

The interesting part of the story occurs in comparing the differences in the two models over time. Table 1.7 shows the differences between the two models across the two scenarios (the difference of differences), which are fairly small for all years. This is likely due to the fact that the demands are very elastic in this model, so even large shocks will cause a small impact on crop prices in these two states. Small price differences correspond to small differences in land allocation and crop supply. Moreover, the calibrated nature of the CC-PE model ensures that the two are reasonably close together. While this ensures that the results are comparable, it may understate the real differences between the two approaches. A modest price increase is sufficient to reallocate corn from other uses to ethanol use, based on this representation. While Table 1.9 shows the gross changes between the two, a closer look at the prices shows a significant difference in the models. In the both models, corn prices increase in accordance with the additional demands of the model. However, the MDP-PE consistently estimates lower prices than the CC-PE model. And, with the exception of 2007 spikes, there is a distinct positive trend in the differences between the two models with regard to price. That is, the prices in the CC-PE model rise at a relatively fast rate, while the MDP-PE prices are slower to respond to the increased demands from the shock. This suggests that the MDP-PE is a more responsive model, reacting more strongly to the issues arising from increased



competition for corn ethanol. That is, a modest change in price will have a greater effect on supply in the MDP-PE model. Looking at acreage in 2008 and 2009 supports this conclusion: the MDP-PE estimated a lower change in prices than the CC-PE, but has more acres planted to each crop in those years.

## **I.10 Conclusion**

The purpose of this research is twofold: can a non-constant land use decision process be incorporated into a structural model with price endogeneity, and, if so, how will it perform against a constant coefficient model? The first question is answered in the affirmative in this case, although there is a need for more refinements in the specification of the structural model. The demand given here is a placeholder, that could be improved, even to the point of introducing a dynamic element to demand. There is also room for improvement as regards the land representation, not least by relaxing the narrow focus on corn and soybean land to incorporate other agricultural land use and other land use categories. More importantly, it was shown that the model may require constraints on the coefficient estimation to ensure that economic theory is followed. Policy has a limited impact through the inclusion of a price floor estimated from marketing loan program impacts. But, it is an advantage of this approach that it permits estimation over small sample sizes that might be more germane to forward-looking decision making as compared to estimates over data extending back to policy and market environments that are no longer relevant. Nevertheless, a possibly quite useful extension would be to extend the model to include more policy effects explicitly, also setting the stage for applied policy analysis.

The second question is answered by comparing a combined MDP-PE with a CC-PE model utilizing a time-averaged set of acreage transition elasticities from the MDP, estimated from 1995-2005. The two models are then shocked with an exogenous demand, increasing net demand to corn by requiring the production of 3.0 billion gallons of ethanol in 2006, and an additional 500 million gallons each year thereafter. The responses of several metrics, including crop prices and area allocated to each crop, were analyzed over 2006-2009. As the two models are calibrated to have similar elasticities at the initial values, a key result is that the MDP-PE shows more price response in the case with the shock. The MDP-PE is shown to allow for increasing responsiveness to changes in prices through estimating consistently lower prices than the CC-PE, while the CC-PE responsiveness tends to be lower path despite changing conditions in this case. This result suggests that the MDP approach might be a means of taking into account dynamic effects, such as rotation or crop-specific human or physical capital, that could affect crop supply elasticities.

## I.11 References

- Adams, G. 1994. "Impact Multipliers of the U.S. Crops Sector: A Focus on the Effects of Commodity Interaction." PhD dissertation, University of Missouri.
- Adelman, I.G. 1958. "A Stochastic Analysis of the Size Distribution of Firms." *Journal of the American Statistical Association* 53: pp. 893- 904.
- Ahn, S., A.J. Plantinga, and R.J. Alig. 2000. "Predicting Future Forestland Area: A Comparison of Econometric Approaches." *Forest Science* 46(3): pp. 363-376.
- Blackmer, A.M., R.D. Voss, and A. P. Mallarino. 1997. "Nitrogen Fertilizer Recommendations for Corn in Iowa." Iowa State University Agr. Ext. Serv. Pm-1714, May.
- Burnham, B.O. 1973. "Markov Intertemporal Land Use Simulation Model." *Southern Journal of Agricultural Economics* 5(1): pp. 253-258.
- De La Torre Ugarte, D.G., M.E. Walsh, H. Shapouri, and S.P. Slinsky. 2003. "The Economic Impacts of Bioenergy Crop Production on U.S. Agriculture." Washington DC: U.S. Department of Agriculture, OEPNU Ag. Econ. Rep. 816, February.
- Devadoss, S. and J. Bayham. 2010. "Contributions of U.S. Crop Subsidies to Biofuel and Related Markets." *Journal of Agricultural and Applied Economics*, 42 (4): pp. 743–756
- Dicks, M.R., J. Campiche, D. De La Torre Ugarte, C. Hellwinkel, H.L. Bryant, and J.W. Richardson. 2009. "Land Use Implications of Expanding Biofuel Demand." *Journal of Agricultural and Applied Economics*, 41 (2): pp. 435–453
- Donner, S.D., and C.J. Kucharik. 2008. "Corn-Based Ethanol Production Compromises Goal of Reducing Nitrogen Export by the Mississippi River." *Proceedings of the National Academy of Sciences of the USA* 105(11): pp. 4513–4518.
- Golan, A., G. Judge, and D. Miller. 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York NY: John Wiley and Sons.
- Golan, A., and S. Vogel. 2000. "Estimation of Non-Stationary Social Accounting Matrix Coefficients with Supply Side Information." *Economic Systems Research* 12(4): pp. 447-471.
- Hallberg, M. 1969. "Projecting the Size Distribution of Agricultural Firms: An Application of a Markov Process with Non-Stationary Transition Probabilities." *American Journal of Agricultural Economics* 51: pp. 289-302.
- Hardie, I.W., and P.J. Parks. 1997. "Land Use with Heterogeneous Land Quality: An Application of an Area Base Model." *American Journal of Agricultural Economics* 79(2): pp. 299-310.

- Hartman, K. 2013. "Iowa Ethanol Corn and Co-Products Processing Values." USDA-MO Dept. of Ag. Market News report NW-GR212. 22 November.
- Heady, E. O. 1954. "Simplified Presentation and Logical Aspects of Linear Programming Technique." *Journal of Farm Economics* 36(5): pp. 1035- 1048.
- Hofstrand, D. and W. Edwards. 2003. "Crop Revenue Insurance." Dept. Econ. FM-1853, Iowa State University.
- Jaynes, E.T. 1957. "Information Theory and Statistical Mechanics." *Physical Review* 106(4): pp. 620-630.
- Kelley, A.C. and L.W. Weiss. 1969. "Markov Processes and Economic Analysis: The Case of Migration." *Econometrica* 37(2): pp. 280-297.
- Khanna, M., H. Önal, X. Chen and H. Huang. 2009. "Meeting Biofuels Targets: Implications for Land Use, Greenhouse Gas Emissions and Nitrogen Use in Illinois Emissions and Nitrogen Use in Illinois" in *Transition to a Bioeconomy: Environmental and Rural Development Impacts* ed. Madhu Khanna. Farm Foundation Conference Proceedings.
- Kim, C.S., G. Schaible, and S. Daberkow. 2010. "The Relative Impacts of U.S. Bio-Fuel Policies on Fuel-Energy Markets: A Comparative Static Analysis." *Journal of Agricultural and Applied Economics*, 42 (1): pp. 121–132
- Lubowski, R.N. 2002. "Determinants of Land-Use Transitions in the United States: Econometric Analysis of Changes Among the Major Land-Use Categories." PhD Dissertation, Harvard University.
- Lubowski, R.N., A.J. Plantinga, and R.N. Stavins. 2003. "Determinants of Land Use Change in the United States 1982-1997." Discussion Paper 03-47, Washington DC: Resources for the Future, August.
- . 2008. "What Drives Land-Use Change in the United States? A National Analysis of Landowner Decisions." *Land Economics* 84 (4): pp. 529–550.
- McFadden, D., and K. Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15: pp. 447-470.
- McRae, E. 1977. "Estimation of Time-Varying Markov Processes with Aggregate Data." *Econometrica* 45: pp. 183-98.
- Miller, D.J. and A.J. Plantinga. 1999. "Modeling Land Use Decisions with Aggregate Data." *American Journal of Agricultural Economics* 81(1): pp. 180-194.
- Miller, J.C. and K.H. Coble. 2011. "Incentives Matter: Assessing Biofuel Policies in the South." *Journal of Agricultural and Applied Economics*, 43 (3): pp.413–421.

- Orazem, P. and J. Miranowski. 1994. "A Dynamic Model of Acreage Allocation with General and Crop-Specific Soil Capital." *American Journal of Agricultural Economics* 76: pp. 385-395.
- Paap, R. and H.K. van Dijk. 2003. "Bayes Estimates of Markov Trends in Possibly Cointegrated Series: An Application to U.S. Consumption and Income." *Journal of Business & Economic Statistics* 21: pp. 547-563.
- Public Law (PL) 110–140. 2007. Energy Independence and Security Act of 2007. [frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi? dbname=110\\_cong\\_public\\_laws& docid=f:publ140.110.pdf](http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ140.110.pdf). March 10, 2008.
- Ray, D.E., D.G. De La Torre Ugarte, M.R. Dicks, and K.H. Tiller. 1998. "The POLYSYS Modeling Framework: A Documentation." Staff Paper, Agricultural Policy Analysis Center, Dept. of Agricultural Economics and Rural Sociology, University of Tennessee.
- Shannon, C.E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27: pp. 379-423, 623-656.
- Shumway, C.R. 1983. "Supply, Demand, and Technology in a Multiproduct Industry: Texas Field Crops", *American Journal of Agricultural Economics* 65 (4): pp. 748-760
- Theil, H. 1969. "A Multinomial Extension of the Linear Logit Model." *International Economic Review* 10: pp. 251-9.
- Thomas, A. 2002. "A Dynamic Model of On-Farm Integrated Nitrogen Management." *European Review of Agricultural Economics* 30: pp. 439-460.
- Tompkin, J.R. 1958. "Response of the Farm Production Unit as a Whole to Prices Response of the Farm Production Unit as a Whole to Prices." *Journal of Farm Economics* 40(5): pp. 1115-1128
- US Department of Agriculture, National Agricultural Statistics Service (USDA/NASS). 2012. "Crop Production 2011 Summary." Available at [usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047](http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047).
- Wallender, S. 2013. "While Crop Rotations Are Common, Cover Crops Remain Rare." *Amber Waves*. 4 March.
- Westhoff, P., and S. Brown. 2009. "US Baseline Briefing Book: Projections for agricultural and biofuel markets." MU-FAPRI Report #01-09, University of Missouri.
- Womack, A. 1976. "The U.S. Demand for Corn, Sorghum, Oats, and Barley: An Econometric Analysis." Economic Report 7615, Department of Agriculture and Applied Economics, University of Minnesota, St. Paul.

Wu, J., and K. Segerson. 1995. "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." *American Journal of Agricultural Economics* 77(4): pp. 1033-1047.

## I.12 Appendix: Figures and Tables

Table 1.1: Markovian Land Transition Probabilities in Iowa, 1996-1997

Source: Author's Estimation

1996	1997			Sum
	Corn	Soy	Residual	
Corn	0.572	0.374	0.054	1
Soy	0.491	0.459	0.051	1
Residual	0.035	0.074	0.891	1

Table 1.2: Demand Equation Price Elasticities

Source: Adaptation of Khanna, et al. 2009

Soybean Demand		Corn Demand	
Variable	Elasticity	Variable	Elasticity
U.S. Corn Price	2.22105	U.S. Soy Price	0.74289
U.S. Soy Price	-4.44211	U.S. Corn Price	-1.48579

Table 1.3: Comparison of Net Returns to Acres Planted Elasticities with Measures of Fit (Mean Absolute Errors), Unconstrained MDP Models versus MDP Model Constrained with Coefficient Constraints and/or Non-Uniform Prior Probabilities

Source: Author's Estimation

		Unconstrained		Constrained	
		Corn Ac	Soy Ac	Corn Ac	Soy Ac
Iowa	Corn NR	0.115	-0.188	0.389	-0.140
	Soy NR	-0.071	-0.016	-0.248	0.140
	MAE	0.020	0.032	0.031	0.023
Missouri	Corn NR	-0.003	-0.119	0.657	-0.120
	Soy NR	-0.062	0.210	-0.466	0.211
	MAE	0.045	0.033	0.058	0.033

Table 1.4: Iowa Imposed Non-Uniform Prior Probabilities for Constrained Markovian Model

Source: Adaptation Based on Wallender 2013

t-1	t			
	Corn	Soy	Residual	Sum
Corn	0.7	0.25	0.05	1
Soy	0.6	0.35	0.05	1
Residual	0.05	0.05	0.9	1

Table 1.5: Comparison of Historical Out of Sample Data with MDP-PE and CC-PE Model Base Scenario Estimates, 2006-2009

Source: Author's Estimation and NASS

<b>Out of Sample Data and Estimate Comparisons by Crop and State, 2006 - 2009</b>							
Area	Variable	Units	Historical				
			2006	2007	2008	2009	Average
Iowa	Corn Acres Planted	Mil ac	12.600	14.200	13.300	13.600	13.425
Iowa	Corn Farmgate Price	\$/bu	2.343	3.178	2.716	2.575	2.703
Iowa	Soy Acres Planted	Mil ac	10.150	8.650	9.750	9.600	9.538
Iowa	Soy Farmgate Price	\$/bu	5.014	7.944	6.747	6.886	6.648
Missouri	Corn Acres Planted	Mil ac	2.700	3.450	2.800	3.000	2.988
Missouri	Corn Farmgate Price	\$/bu	2.465	3.263	2.672	2.578	2.745
Missouri	Soy Acres Planted	Mil ac	5.150	4.700	5.200	5.350	5.100
Missouri	Soy Farmgate Price	\$/bu	5.088	8.171	6.693	6.972	6.731
National	Corn Farmgate Price	\$/bu	1.724	2.492	3.383	2.664	2.566
National	Soy Farmgate Price	\$/bu	4.271	5.689	7.999	7.175	6.283
Area	Variable	Units	Markov Decision Process				
			2006	2007	2008	2009	Average
Iowa	Corn Acres Planted	Mil ac	12.411	12.817	12.743	12.758	12.682
Iowa	Corn Farmgate Price	\$/bu	2.137	2.140	2.261	2.238	2.194
Iowa	Soy Acres Planted	Mil ac	10.359	10.244	10.411	10.469	10.371
Iowa	Soy Farmgate Price	\$/bu	5.416	5.390	5.750	5.493	5.512
Missouri	Corn Acres Planted	Mil ac	2.863	2.963	2.888	2.823	2.884
Missouri	Corn Farmgate Price	\$/bu	2.218	2.221	2.337	2.315	2.273
Missouri	Soy Acres Planted	Mil ac	5.013	5.003	5.057	5.070	5.036
Missouri	Soy Farmgate Price	\$/bu	5.465	5.439	5.792	5.541	5.559
National	Corn Farmgate Price	\$/bu	2.458	2.463	2.643	2.609	2.543
National	Soy Farmgate Price	\$/bu	5.468	5.449	5.712	5.524	5.538
Area	Variable	Units	Constant Coefficient				
			2006	2007	2008	2009	Average
Iowa	Corn Acres Planted	Mil ac	12.295	12.768	12.569	12.561	12.548
Iowa	Corn Farmgate Price	\$/bu	2.164	2.154	2.295	2.275	2.222
Iowa	Soy Acres Planted	Mil ac	10.327	10.265	10.314	10.325	10.308
Iowa	Soy Farmgate Price	\$/bu	5.499	5.431	5.854	5.604	5.597
Missouri	Corn Acres Planted	Mil ac	2.714	2.859	2.760	2.701	2.759
Missouri	Corn Farmgate Price	\$/bu	2.244	2.235	2.369	2.349	2.299
Missouri	Soy Acres Planted	Mil ac	4.812	4.867	4.880	4.909	4.867
Missouri	Soy Farmgate Price	\$/bu	5.546	5.479	5.895	5.650	5.643
National	Corn Farmgate Price	\$/bu	2.498	2.484	2.694	2.663	2.585
National	Soy Farmgate Price	\$/bu	5.528	5.479	5.789	5.606	5.601



Table 1.6: Differences between Estimated Markovian and Constant Coefficient Model Base Scenarios, 2006-2009, Constant Coefficient Model Shock Estimates Subtracted from Markov Model Shock Estimates

Source: Author's Estimation

<b>Base Scenario Difference, MDP-PE - CC-PE</b>							
Area	Variable	Units	2006	2007	2008	2009	Average
Iowa	Corn Acres Planted	Mil ac	0.116	0.049	0.174	0.197	0.134
Iowa	Corn Farmgate Price	\$/bu	-0.027	-0.014	-0.034	-0.037	-0.028
Iowa	Soy Acres Planted	Mil ac	0.032	-0.021	0.097	0.144	0.063
Iowa	Soy Farmgate Price	\$/bu	-0.083	-0.041	-0.104	-0.111	-0.085
Missouri	Corn Acres Planted	Mil ac	0.149	0.104	0.128	0.122	0.126
Missouri	Corn Farmgate Price	\$/bu	-0.026	-0.014	-0.032	-0.034	-0.027
Missouri	Soy Acres Planted	Mil ac	0.201	0.136	0.177	0.161	0.169
Missouri	Soy Farmgate Price	\$/bu	-0.081	-0.040	-0.103	-0.109	-0.083
National	Corn Farmgate Price	\$/bu	-0.040	-0.021	-0.051	-0.054	-0.042
National	Soy Farmgate Price	\$/bu	-0.060	-0.030	-0.077	-0.082	-0.062

Table 1.7: Markov Model Scenario Differences, 3.0 Billion Gallon Shock Minus Baseline Scenario, 2006-2009

Source: Author's Estimation

<b>MDP-PE Shock Scenario Difference</b>							
Area	Variable	Units	2006	2007	2008	2009	Average
Iowa	Corn Acres Planted	Mil ac	0.000	0.443	0.626	0.898	0.492
Iowa	Corn Farmgate Price	\$/bu	0.690	0.680	0.854	0.849	0.768
Iowa	Soy Acres Planted	Mil ac	0.000	-0.034	0.103	0.191	0.065
Iowa	Soy Farmgate Price	\$/bu	1.428	1.397	1.662	1.599	1.522
Missouri	Corn Acres Planted	Mil ac	0.000	0.122	0.172	0.242	0.134
Missouri	Corn Farmgate Price	\$/bu	0.658	0.648	0.813	0.808	0.732
Missouri	Soy Acres Planted	Mil ac	0.000	0.088	0.166	0.219	0.118
Missouri	Soy Farmgate Price	\$/bu	1.400	1.370	1.630	1.567	1.492
National	Corn Farmgate Price	\$/bu	1.030	1.014	1.274	1.266	1.146
National	Soy Farmgate Price	\$/bu	1.045	1.022	1.217	1.170	1.114
National	Ethanol Corn Net Demand	Mil bu	-758.9	885.4	1011.9	1138.3	948.6
Regional	Non-Ethanol Corn Demand	Mil bu	-758.9	-796.7	-886.0	-947.1	-847.2
Regional	Corn Production	Mil bu	0.000	88.7	125.9	191.3	101.5

Table 1.8: Constant Coefficient Model Scenario Differences, 3.0 Billion Gallon Shock Minus Baseline Scenario, 2006-2009  
Source: Author's Estimation

CC-PE Shock Scenario Difference							
Area	Variable	Units	2006	2007	2008	2009	Average
Iowa	Corn Acres Planted	Mil ac	0.000	0.454	0.449	0.645	0.387
Iowa	Corn Farmgate Price	\$/bu	0.718	0.692	0.960	0.985	0.839
Iowa	Soy Acres Planted	Mil ac	0.000	-0.031	-0.035	-0.067	-0.033
Iowa	Soy Farmgate Price	\$/bu	1.475	1.416	1.874	1.880	1.661
Missouri	Corn Acres Planted	Mil ac	0.000	0.116	0.113	0.168	0.099
Missouri	Corn Farmgate Price	\$/bu	0.684	0.659	0.915	0.939	0.799
Missouri	Soy Acres Planted	Mil ac	0.000	0.093	0.088	0.104	0.071
Missouri	Soy Farmgate Price	\$/bu	1.446	1.389	1.836	1.842	1.628
National	Corn Farmgate Price	\$/bu	1.071	1.033	1.432	1.471	1.252
National	Soy Farmgate Price	\$/bu	1.080	1.036	1.371	1.375	1.216
National	Ethanol Corn Net Demand	Mil bu	758.9	885.4	1011.9	1138.3	948.6
Regional	Non-Ethanol Corn Demand	Mil bu	-758.9	-795.5	-922.8	-1001.6	-869.7
Regional	Corn Production	Mil bu	0.000	89.839	89.093	136.7	78.9

Table 1.9: Difference of Differences Across the Markov and Constant Coefficient Models, Constant Coefficient Model Shock Difference Estimates Subtracted from Markov Model Shock Difference Estimates  
Source: Author's Estimation

Difference of Markov and Constant Coefficient Scenario Differences							
Area	Variable	Units	2006	2007	2008	2009	Average
Iowa	Corn Acres Planted	Mil ac	0.000	-0.011	0.177	0.253	0.105
Iowa	Corn Farmgate Price	\$/bu	-0.028	-0.012	-0.106	-0.136	-0.071
Iowa	Soy Acres Planted	Mil ac	0.000	-0.003	0.138	0.258	0.098
Iowa	Soy Farmgate Price	\$/bu	-0.047	-0.019	-0.212	-0.281	-0.140
Missouri	Corn Acres Planted	Mil ac	0.000	0.006	0.059	0.074	0.035
Missouri	Corn Farmgate Price	\$/bu	-0.026	-0.011	-0.102	-0.131	-0.068
Missouri	Soy Acres Planted	Mil ac	0.000	-0.005	0.078	0.115	0.047
Missouri	Soy Farmgate Price	\$/bu	-0.046	-0.019	-0.206	-0.275	-0.137
National	Corn Farmgate Price	\$/bu	-0.041	-0.019	-0.158	-0.205	-0.106
National	Soy Farmgate Price	\$/bu	-0.035	-0.014	-0.154	-0.205	-0.102
Regional	Corn Production	mil bu	0.000	-1.173	36.803	54.575	22.551

Figure 1.1: Illustration of the MDP-SPE Model Framework  
 Source: Author's Representation

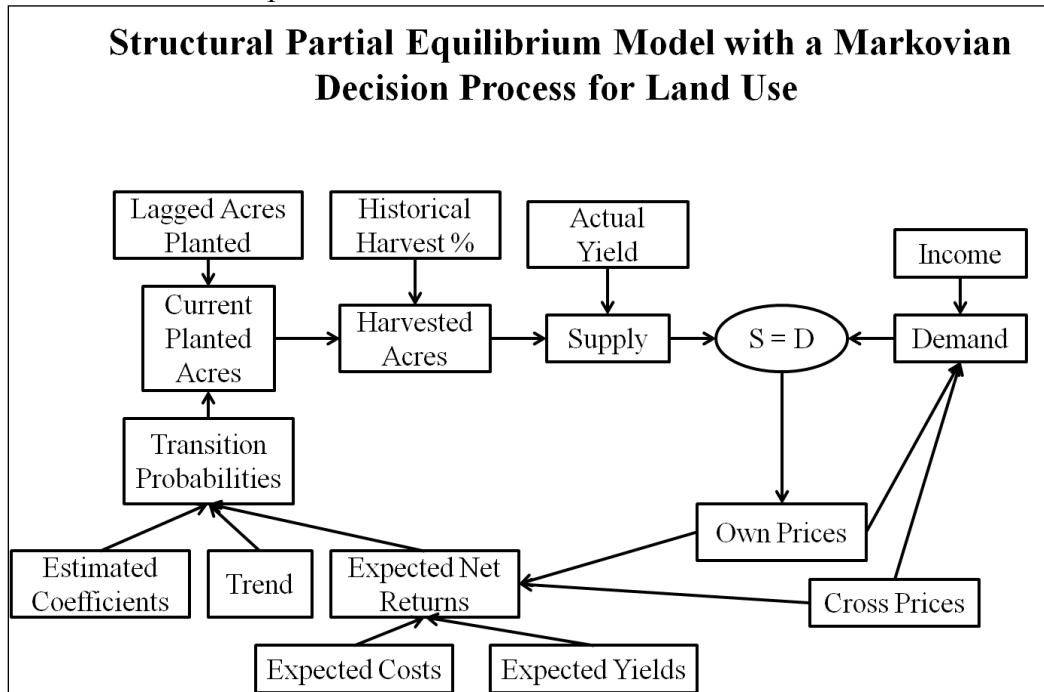


Figure 1.2: Markov Decision Process Model and Constant Coefficient Model Corn and Soybean Acres Planted Comparison to Historical Data, Iowa, 1996 – 2009  
 Source: Author's Estimates and NASS

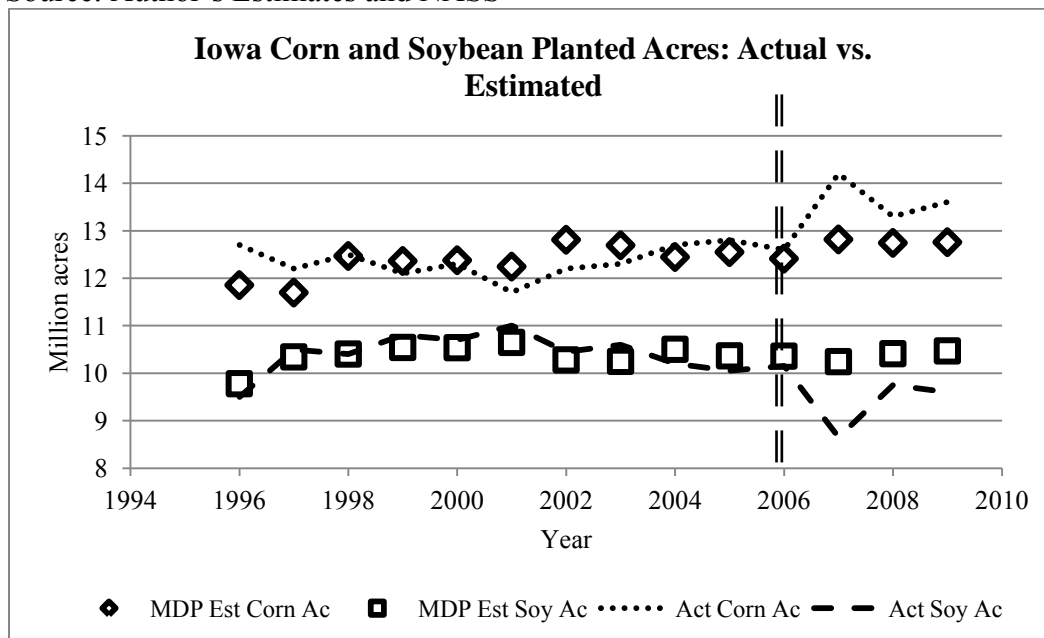
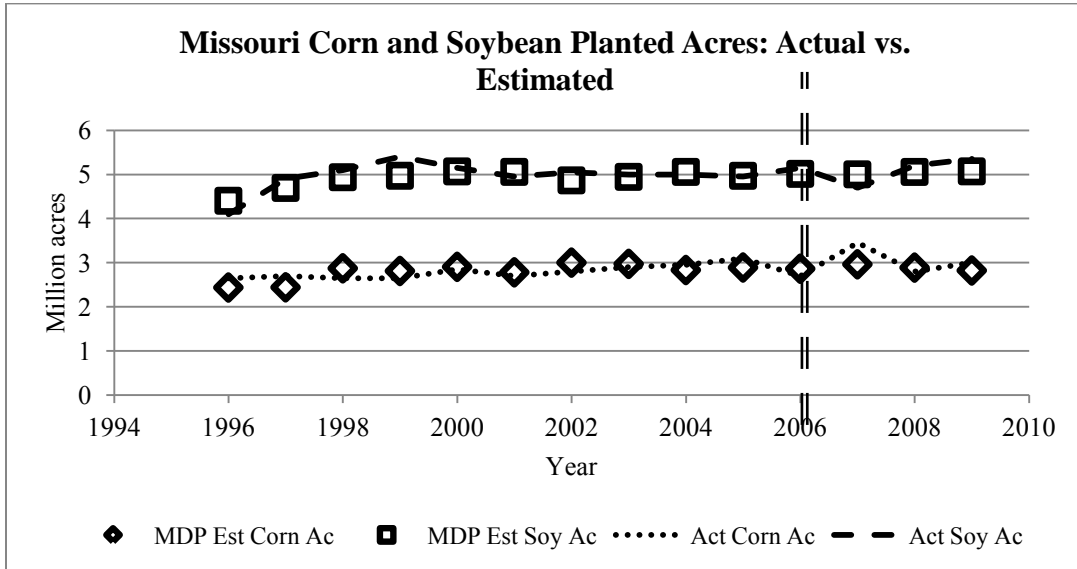


Figure 1.3: Markov Decision Process Model and Constant Coefficient Model Corn and Soybean Acres Planted Comparison to Historical Data, Missouri, 1996 – 2009  
 Source: Author's Estimates and NASS



## **II. An Economic Examination of Midwestern Warm Season Grass Area Using Satellite Imaging Data**

### **II.1 Introduction**

In recent years, the United States has increased the focus it places on biofuel policy as a part of its overall energy policy. Specifically, the Energy Independence and Security Act (EISA) of 2007 includes mandates for the use of several categories of biofuel, part of which can be met using conventional (corn starch based) and cellulosic ethanols over the next decade. These mandates (which may or may not be waived at the discretion of the Environmental Protection Agency) could significantly change the manner in which U.S. land is allocated for crop production for the next decade.

One source of cellulosic biofuels is from warm-season grasses (WSG). WSG are perennial grasses, normally planted in the spring, and harvested multiple times over the course of the year. Notable WSG include switchgrass (*Panicum virgatum L.*), miscanthus (*Miscanthus giganteus*), Indian grass (*Sorghastrum nutans*), and big bluestem (*Andropogon gerardii*). These grasses grow naturally in the Midwestern United States, and have been considered as a means of supplementing, and to some degree possibly replacing, corn as an fuel ethanol feedstock. As a commodity, WSG may compete with traditional crops for land and capital. Understanding this competition would be useful to policymakers in crafting legislation. However, the nature of this competition is not initially clear.

WSG data are somewhat limited. While studies have been done to ascertain yields and budgets, these studies are relatively few and focused on small areas, chiefly owing to a lack of data. The National Agricultural Statistics Service (NASS) does not track WSG

production as it does other crops. Normally, land use studies have the amount of land dedicated to a particular crop explained by variables such as own and cross prices, yields, fertilizer costs, government policies, and other variables that influence a farmer's planting decision. Without the data to be explained, however, such studies cannot be conducted on a large scale.

However, some such data have recently become available. Using the National Aeronautics and Space Administration's (NASA) Moderate Resolution Imaging Spectroradiometer (MODIS) satellites, land cover images were obtained detailing acres where WSG were present in the Midwestern U.S. These data, collected at the county level from 2000-2009 (see Wang et. al 2011), can serve as the left-hand-side variable in an econometric analysis to estimate the relationship between WSG and other crops, based on the aforementioned explanatory variables.

This research performs such an analysis, examining the historical growth of WSG in response to own and cross price, yield, and cost effects. However, the data are still limited, and the structure of the market might not reflect what would happen if cellulosic biofuel takes off, given the relatively new nature of WSG being used as a feedstock. So, a more frequently used method of econometric analysis, such as Ordinary Least Squares (OLS), may not be appropriate. A method which is robust with limited data, or is able to handle ill-posed problems, may provide a better understanding of the relationship between WSG and traditional crops. To wit, this research will assume land use patterns follow a first-order Markov Decision Process (MDP). It is a dynamic process in which the current land allocation is a function of the last period's allocation and a set of non-

stationary transition probabilities. These probabilities are estimated based on data which affect the decision process.

Thus, the question this research intends to answer is “What relationships, if any, exist between warm season grasses and major traditional crops in the Midwestern US?” To answer this question, this research estimates WSG, corn, soy, and wheat area using an MDP.

## **II.2 Prior Literature**

With the increasing importance of biofuels in the American agricultural economy, literature concerning such impacts on land use and other areas has increased in frequency. Devadoss and Bayham (2010) examine the effects of reducing crop subsidies on biofuel production. Dicks et al. (2009) examine the amounts of various feedstocks necessary to satisfy the EISA mandate. Miller and Coble (2011) study different possible policy options available to lawmakers and the impact of those proposals on the Southeastern U.S. Kim, Schaible, and Daberkow (2010) focus on the impacts of U.S. biofuel policy on international energy markets. While this list of biofuels related economic literature is far from exhaustive, it makes clear that there are a great many questions to answer about the changing bioenergy economy. This research attempts to answer a couple of those questions.

Earlier, the availability of local or regional studies for WSGs is mentioned. These budgets, while unable to encompass the entire area of study, are important because they show different techniques and costs at a local level. These budgets often compare hay with switchgrass, because the production processes are so similar. Such budgets have been prepared by Smith (2009), University of Tennessee Extension (2009), Carpenter and

Brees (2010), Holman et al. (2013), Duffy and Nanhou (2001), Helsel and Alvarez (2012), and Brechbill, Tyner, and Ileleji (2011). However, two papers in particular have very wide study areas, and therefore become of particular interest: Perrin et al. (2008) covers fields in North Dakota, South Dakota, and Nebraska. Also, Khanna, Jain, and Oliver (2011) cover Minnesota, Wisconsin, Michigan, Iowa, Missouri, Illinois, Indiana, and Ohio. These two studies cover the bulk of available area, and are used as the primary sources of cost data for this study.

The earliest use of a MDP to estimate changes in land use is found in Burnham (1973). The author assumes that land use changes in the Southern Mississippi Alluvial Valley can be estimated by a stationary first order Markov chain. However, he states in the footnotes that assuming stationary transition probabilities may be too restrictive for a land use change model. Burnham's conclusion is supported by Hallberg (1969), who studies frozen dairy products in Pennsylvania using a non-stationary MDP. Hallberg uses multiple regression techniques to test the hypothesis that non-stationary Markov transition probabilities have better predictive capability than stationary probabilities. MDPs are used in far more than land use, however. Again, Hallberg's study is on firm size in the dairy industry, not land use. Other studies using MDPs for econometric analysis include Adelman's analysis of the distribution of firm sizes in the iron and steel industry (1958); Paap and Van Dijk's analysis of income and consumption in the U.S. (2003); Kelley and Weiss' study of population migration based on wage differentiation (1969); Miller and Plantinga's paper analyzing land use changes in Iowa (1999); and Lubowski, Plantinga, and Stavins' use of a nested logit model to analyze national non-federal land use (2008).



As stated earlier, other models are utilized in economics to determine land use allocation. One such method is an iterative linear program (LP). In an LP, a linear production function is optimized subject to constraints such as total acreage availability. Heady (1954) provides a discussion of the logic and advantages of using linear programs in agricultural econometrics. Tompkin (1958) uses linear programming to determine the optimal combination of production activities on a livestock farm. A different type of LP model is used by De La Torre Ugarte et al. (2003) and Ray et al. (1998) in the implementation of the POLYSYS agricultural policy model. However, LP models may result in corner solutions, so transition constraints may be imposed by the modeler to prevent such scenarios.

Another model used to estimate land use shares and transition probabilities is the multinomial logit model (MNL) (Theil 1969). The MNL is used by McRae (1977), among others. The probabilities associated with a change in state are estimated using a logistical form. This functions similarly to a Markov chain, but disregards the prior state, estimating changes in land use solely as a function of exogenous data. Such use can be seen in Wu and Segerson (1995); Hardie and Parks (1997); and Ahn, Plantinga, and Alig (2000), for instance. In addition, Lubowski, Plantinga, and Stavins (2003, 2008) note that the Independence from Irrelevant Alternatives (IIA) property of MNLs may preclude otherwise optimal choice behaviors. They, in addition to Lubowski (2002), use a model known as a nested logit model (NLM) to address this shortcoming of the MNL. The NLM separates decision states into subgroups or “nests” of similar qualities, differentiating them based on degree of substitutability. The nests Lubowski, Plantinga, and Stavins use include urban, non-farm (comprised of forest and range land), and farm

(comprised of cropland, Conservation Reserve Program land, and pasture land). The advantage of the NLM is that it imposes IIA within nests, but not across nests, relaxing the choice restrictions. However, because the nests of a NLM are based on substitutability, it may be less efficient at explaining land use change between crops. The mixed logit model (MLM) is another method of relaxing the IIA restriction. The MLM is given by McFadden and Train (2000), and takes the choice specific variables of conditional logit and the choice-independent variables from the MNL to create a mixed model in which additional choices change the relative probabilities of the existing states.

However, the MLM still requires a problem that is well posed: i.e. the observations available exceed the number of unknowns. In the case of the introduction of a new crop for cellulosic biofuel feedstocks, the known data are very limited, and the problem may be ill-posed. Golan, Judge, and Miller (1996) discuss the use of general maximum entropy to address the issue of an ill-posed problem. From this discussion and other existing land use literature, a model that estimates a set of coefficients directly linking explanatory state variables (both state dependent and independent to avoid IIA) with decision variables, as well as incorporating data from the state of the land in the prior period may be derived to provide a more robust analysis. One such model that fits those criteria is the MDP model.

### **II.3 Markovian Land Use Decision Process**

This research examines the dynamics of land use changes by explaining the planting decisions of farmers in the area of study. Following Ahn, Plantinga, and Alig (2000), each farmer in region  $i$  ( $i = 1, \dots, I$ ) is assumed to plant a sequence of crops on land that he manages that maximizes the present discounted value of expected net returns

$$(2.1) \quad \max_{\delta_{j,k,t}} \sum_{t=0}^{\infty} \tau^t \cdot E[NR(\delta_{j,k,t}, \mathbf{X}_{h,i,t}) + \varepsilon_t(\delta_{j,k,t})]$$

where  $\tau$  is a constant discount factor,  $\delta_{j,k,t}$  represents a decision to allocate land currently used for crop (discrete state variable)  $j$  in time  $t-1$  to crop  $k$  in time  $t$ ,  $\mathbf{X}_{h,i,t}$  represents a  $T \times H$  matrix of observable data, including expected crop net returns and a trend variable, and  $\varepsilon_t$  represents unobserved variables. Beyond the limit of a farmer making decisions only for the land he owns or controls, this maximization is unconstrained. Because of the relative difficulty involved with obtaining individual farmer planting decisions, a model which utilizes aggregated data is desirable. Therefore, aggregated acres planted are used to represent the sum of all planting decisions (Golan, Judge, and Miller 1996). However, there exist idiosyncratic or agent-specific components of a farmer's planting decision, which are unobservable to the econometrician. These components are represented by the unobservable state variable  $\varepsilon$ . Over time, a farmer's planting decisions are assumed to follow a first order MDP. The acreage planted to a given crop is a function of the crop planted in the prior period and a non-stationary transition probability, noted  $\pi_{i,j,k,t}$ . That is, there is a non-observable  $J \times K$  (where  $J=K$ ) matrix of probabilities for each state and time period, transitioning from crop  $j$  to crop  $k$  from period  $t-1$  to period  $t$ .

To illustrate, Table 2.1 shows a set of probabilities for land use transition from 2001-2002 as might be estimated from the model explained below. If an acre of land was planted with corn in 2001, it has approximately a 26% chance of staying in corn in 2002, a 25% chance of being planted with soybeans, a 3% chance of being planted with wheat, a 6% chance of being planted with WSG, and a 40% chance of being used in another category. These numbers are fictional and only present to provide a visual example of the

nature of Markovian transition probabilities; they are not results. Markovian transition probabilities row-sum to one and are hypothesized to be affected by explanatory variables including crop prices. Through this interaction, the variables are explanatory of changes in crop land planting patterns.

The model is derived using maximum entropy (ME), following Golan and Vogel (2000); Golan, Judge, and Miller (1996); and Miller and Plantinga (1999). The ME method for estimating the model of Markov transition probabilities is set forth by Jaynes (1957). The objective of the ME method is to select the probabilities that use the least information (fewest assumptions) while still satisfying the constraints. This, in turn, assumes the farmer has the greatest amount of choice possible. Shannon's (1948) entropy measure is used to measure the amount of information needed to estimate the coefficients (Miller and Plantinga 1999, Golan and Vogel 2000). The primal objective function determining the optimal transition probabilities is

$$(2.2) \quad \begin{aligned} \max_{\pi_{i,j,k,t}} S_i &= \sum_{t=1}^T \sum_{j=1}^J \sum_{k=1}^K \pi_{i,j,k,t} \cdot \ln(\pi_{i,j,k,t}) \\ \text{s.t.} \quad \sum_{t=1}^T \mathbf{X}'_{i,t} \left( y_{i,k,t} - \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t} \right) &= \bar{0} \end{aligned}$$

where  $\bar{0}$  is an  $H$  size vector of zeroes and  $y_{i,k,t}$  is the share of acreage planted to crop  $k$  in area  $i$  during time  $t$ . When applied to the estimating equations (the set of constraints), the solution to the problem takes the form

$$(2.3) \quad \hat{\pi}_{i,j,k,t} = \frac{q_{i,j,k,t} \cdot \exp\left(y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \hat{\lambda}_{h,i,k}\right)}{\sum_{k=1}^K q_{i,j,k,t} \cdot \exp\left(y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \hat{\lambda}_{h,i,k}\right)},$$

where  $\hat{\lambda}_{h,i,k}$  is the optimal Lagrangian multiplier associated with explanatory variable  $h$  and crop  $k$ , and  $q_{i,j,k,t}$  is a conditional probability, which may be adjusted by the analyst to represent information from before the planting decision that may bias the farmer's decision. Here, priors are not uniform, but the exact structure of the priors will be discussed later. By focusing on the dual of the primal problem, an unconstrained equation to determine the optimal multipliers takes the form of

$$(2.4) \quad \max M_i(\bar{\lambda}) = \sum_{t=1}^T \sum_{h=1}^H \sum_{k=1}^K y_{i,k,t} x_{h,i,t} \lambda_{h,i,k} - \sum_{t=1}^T \sum_{j=1}^K \ln \left[ \sum_{k=1}^K q_{i,j,k,t} \exp \left( y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \lambda_{h,i,k} \right) \right]$$

One  $k$  category is kept as a residual, with its multipliers assumed to be zero. Because of the additive nature of the land categories, the residual solution is implicit when all other categories are estimated. It should be noted here that the residual category was shock tested after the coefficients were estimated, and was found to move as expected with regard to changes in explanatory data: increases in all net returns categories caused a marked decrease in residual category land, while increasing all other land categories. After the multipliers are estimated, they are used with the other variables to determine the transition probabilities in (2.3). The transition probabilities are then applied to the prior year's crop plantings to determine current plantings:

$$(2.5) \quad \hat{y}_{i,k,t} = \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t}$$

With the acreage shares estimated, elasticities can be calculated to show the effect of a one percent change in the explanatory variables on the transition probabilities. (Miller and Plantinga 1999)

$$(2.6) \quad \omega_{h,i,j,k,t} = x_{h,i,t} \cdot y_{i,j,t-1} \left[ \hat{\lambda}_{h,i,k} - \sum_{k=1}^K \hat{\lambda}_{h,i,k} \cdot \pi_{i,j,k,t} \right]$$

Using this equation, one can derive the acreage elasticities, measuring the change in the acreage allocation from a one percent change in the explanatory variables.

$$(2.7) \quad \kappa_{h,i,k,t} = \hat{y}_{i,k,t}^{-1} \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t} \cdot \omega_{h,i,j,k,t}$$

Using the appendix in Miller and Plantinga (1999) as a guide, the covariance matrix of the coefficients ( $\hat{\lambda}_{h,i,k}$ ) conditional on the explanatory variables can be estimated by

$$(2.8) \quad \left( (\mathbf{X} \otimes \mathbf{I}_{(K-1)^2})' \boldsymbol{\Sigma}_i (\mathbf{X} \otimes \mathbf{I}_{(K-1)^2}) \right)^{-1}$$

such that  $\mathbf{I}$  is the identity matrix, and  $\boldsymbol{\Sigma}_i$  is a  $T(K-1)^2 \times T(K-1)^2$  matrix where the diagonal elements are  $\pi_{i,j,k,t}(1 - \pi_{i,j,k,t})$  and the off diagonal elements are  $\pi_{i,j,k,t}(-\pi_{i,j,k,t})$ .

With the MDP model defined, the focus shifts to defining the data that will be used in the model.

Table 2.2<sup>1</sup> shows historical averages of the shares and expected net returns by crop. Shares of and net returns to WSG are relatively small, historically. Corn and soybeans are the most prevalent crops in the area of study, and those with the highest net returns, presumably in part because they tend to be grown on the land with better than average growing conditions. Changes to net returns per acre in WSG may need to be large (such as a 50% or 100% increase) to cause a shift in acreage away from corn and soybeans in such areas. Wheat also has a relatively small area, but the net returns are higher than those of WSG. This suggests that WSG may compete more easily with wheat,

---

<sup>1</sup> Table 2 does not include Indiana, Michigan, Northeastern Minnesota, or Ohio, as explained in the next section.

which is to say that a more modest change in WSG returns might cause wheat acreage to decrease. Wheat area may seem too low to the informed reader. However, the acreage share estimates for corn, soy, and wheat were taken directly from NASS; only WSG land cover area is taken from MODIS. Because this research is dealing with shares of total land, it may skew perception. To double check, random shares were taken from the data and multiplied by total CRD area to obtain gross acreage estimates. The resulting estimates matched NASS data to within one acre, which can be assumed as rounding error.<sup>2</sup>

#### **II.4 Data: WSG Returns**

This research is chiefly concerned with the inclusion of a new data source in the estimation of land allocated to the production of WSG: NASA satellite images. The satellite image source has been previously discussed, but it is important to discuss the geographic nature of the data. As shown in Figure 2.1, WSG land cover has been prevalent in areas with large areas of pastureland, particularly areas like Eastern Nebraska and North Dakota. This is important because it informs the discussion of a proxy price for WSG. There is no historical price for WSG, so a proxy must be chosen. Among options considered was assuming a biomass price and conversion rate based on prior literature, such as those found in Brechbill, Tyner, and Ileleji (2011); De La Torre Ugarte et al. (2003); or Khanna, Shungana, and Clifton-Brown (2008), but this was seen as problematic because of the lack of variation and the inability to provide a substantive motivation for variation. Another idea was to use the price of hay, but this datum was difficult to obtain at a county level. Because of the strong link between WSG production and pastureland, the county-level pastureland rental rate (PRR) collected by NASS was

---

<sup>2</sup> Table 6 shows the model estimates for gross acreage, which may be more in line with expected amounts.

chosen as the proxy for WSG production net returns<sup>3</sup>. The PRR, while considered a good proxy for historical WSG returns, is not collected by NASS for the eastern states in the area of study, namely Indiana, Ohio, and Michigan, along with one Crop Reporting District<sup>4</sup> (CRD) in Minnesota. Given that area allocated to WSG in these regions is relatively low, they were excluded from the area of study. Further, NASS only reports PRR at the county level from 2008-2012, but at state level for all years in the study. To spread the PRR across the missing years, the average absolute difference between county and state PRR were taken as a measure of basis, which is assumed to be constant over time. This difference was applied to the historical state rates to get historical county rates. This provided both the spatial and temporal variation desired for this research, and is used to estimate the coefficients mentioned above for WSG returns.

## **II.5 Data: Other Returns**

Other explanatory data in the model include expected net returns for corn, soy, and wheat, as well as a trend variable and an intercept. Expected net returns for established crops are estimated using county-level trend yields, variable costs of production, and a proxy for the forward contract price at the county elevator. Expected yields are trend yields estimated from county level data taken from NASS. Costs of production are taken from ERS, and assigned to counties based on which Farm Resource Region they fall in. To obtain county-level futures prices, prices reported at grain elevators are interpolated at the county level.<sup>5</sup> These prices are compared to national prices reported from NASS, and the average difference over time is calculated to act as

---

<sup>3</sup> The pastureland rental rate, as well as all prices, are in real terms.

<sup>4</sup> Also known as Agricultural Statistical Districts (ASDs).

<sup>5</sup> The author wishes to thank Tim Matisziw of the MU Department of Geography for his work in interpolating the grain elevator prices.



an estimate of basis. This connects easily obtainable national level data to county level data. To link futures prices to the county level prices, the Chicago Mercantile Exchange (CME) futures price was estimated as a function of the lagged national marketing year crop price. This created an “intertemporal basis” estimate, allowing national futures prices to be disaggregated down to the county level. Daily futures prices were obtained from CME. Following the procedure used by crop insurance (Hofstrand and Edwards 2003), the average of daily settlement prices in February were used to calculate the futures price used. Corn and wheat prices were based on December contracts, while soybean futures prices were based on November contracts.

## **II.6 Data: Acreage Shares**

Estimates of acreage allocated to each use is one of the goals of this study. As such, acreage data are necessary. County level acreage estimates are taken from NASS for corn, soy, and wheat. WSG acreage is taken from the NASA MODIS images mentioned before and aggregated to county level. All other land in a county is aggregated into a single residual category. All other land, as opposed to all agricultural land or all cropland, is chosen for two reasons. 1) The author’s prior experience with the MDP has shown that small residual categories can be problematic in estimation, so the largest residual category possible is desirable. 2) WSG can be produced on lands other than those traditionally allocated to production agriculture, including CRP land. The possibility of production in these other areas needs to be taken into account. This provides an additional benefit in that it keeps total area constant over time, removing the need to estimate and possibly introducing further error. Due to the probabilistic nature of the MDP, shares of total area are used instead of gross acreage.

Using mixed datasets can cause issues in estimation. However, this research seeks to use as much publicly available data as possible, and NASS estimates are generally accepted. However, as Wang et al. (2011) note, there may be an issue with the estimation of WSG area with regards to spring wheat. The system used to classify the data could cause spring wheat to be classified as WSG. While the author was not able to independently confirm the existence of misclassification errors, it is important to discuss the errors' possible implications. First, wheat would be double counted in such an instance, which would bias the share estimates which serve as the left-hand-side variable for this research. WSG shares would be too high, since non-WSG area is included in the estimate, which would bias all other share estimates downward. Second, with area misclassified, responses to changes in explanatory data may not follow expectations. That is, if area classified as WSG were actually wheat, it would respond positively to changes in wheat net returns, as opposed to negatively, as would be expected for competitive WSG. While these issues are recognized, the area classified as WSG is assumed to be WSG in the absence of a method to independently confirm or correct for such errors.

## **II.7 Data: Prior Probabilities**

As mentioned earlier,  $q_{j,k,t}$  is a set of prior probabilities which represents a bias in the decision process (2.1) imposed by the econometrician to cause the probabilities estimated in the model to reflect the farmer's decision process more accurately. Table 2.3 shows the prior probabilities implemented in the model. These numbers were arrived at by taking a minimum set of priors (all set at .05) and modifying them based on certain assumptions about production agriculture. 1) Corn/soy rotation is strongly followed in the Midwest, increasing corn/soy and soy/corn probabilities. 2) Soy/Wheat double cropping

and soy/wheat rotations occur in some areas of study, increasing soy/wheat and wheat/soy probabilities. Double cropping does not appear in the data explicitly, but instead as concurrent acres of soybean and wheat in a single year. The approach here increases the potential for soybean area leading to more wheat area and wheat area leading to more soybean area. These greater probabilities would make sense if in continuous rotation, but overlooks the potential that soybean and wheat that are double cropped in one year will be rotated to another crop in the next year. The increase in probabilities to take soybean and wheat rotation into account is smaller than the increase for corn and soybean rotation, corresponding to the expectation about which of these planting parameters comprises a greater share of the relevant crop areas. 3) Residual area, because it includes urban and non-agricultural (marginal) land, is likely to remain in that use once area is allocated to it, increasing residual/residual probabilities. 4) Because WSG requires an establishment period of three to five years and is a perennial, area allocated to WSG is more likely to remain in WSG, increasing WSG/WSG probabilities.

## **II.8 Spatial Correlation Issues**

At county level, working with a dataset of this size can present some challenges. For the model to solve in a reasonable amount of time, each county was coded to solve in autarky. However, this presents issues with likely correlation of errors across counties due to factors such as weather. A higher level of aggregation, such as national data, can address such issues, but heterogeneity across land types is lost at such a level. One possible medium is solving with acreage aggregated and explanatory data averaged at CRD level. It should be noted that, in this manner, each CRD solves in autarky, but, this is considered an acceptable assumption, because the correlations across CRDs are likely

much weaker than the correlations across counties. There are external factors that can affect wide areas, such as weather patterns in general and particularly large-scale droughts, that suggest correlation among CRDs, but these interactions in error terms are omitted here.

To compare the two aggregation levels, corn own net return elasticities, wheat acreage-PRR elasticities, and average error were calculated after the estimation of the coefficients and compared. A histogram of county level and CRD level corn net returns to corn area planted elasticities can be found in Figures 2.2 and 2.3. The shapes of the distributions are very similar, though the county level has a slightly higher mean, indicating a greater level of response. Figures 2.4 and 2.5 show the returns of wheat acreage to PRR, the WSG net return proxy. The shapes of the two histograms are likewise similar, but the county level has a greater percentage of counties and years with a positive elasticity, which does not follow expectations: given the tendency of WSG to grow in poorer quality soil, it is expected to compete with other crops that grow in similar conditions, namely wheat. It should be noted here that the magnitude of the means reported in the histograms cannot be considered a true measure of the effects, because, to allow for proper scaling and visual representation of the data, any elasticities greater than three or less than negative three were forced to equal to those values, respectively<sup>6</sup>. But, even taking that into account, one can see that wheat area has a very strong reaction to changes in PRR. These two sets of figures show that the CRD level of aggregation should perform in much the same manner as the county level in autarky. But, the question remains as to the relative level of performance between the two options. To address this

---

<sup>6</sup> This change in elasticities is purely for cosmetic purposes, and does not change the values of the coefficients.

concern, the average errors were estimated for the acreage shares to act as an indicator of fit:

$$(2.9) \quad MAE_k = \frac{1}{TI} \sum_{t=1}^T \sum_{i=1}^I \frac{abs(y_{i,t,k} - \hat{y}_{i,t,k})}{y_{i,t,k}}$$

MAE is the percent mean absolute error: the absolute difference of the estimated acreage and the actual acreage over time for each crop, normalized by the actual acreage planted to that crop, averaged over time. As it measures absolute deviation from historical data, a lower number indicates a better fit. Because  $y$  represents a share, as opposed to actual acreage, this estimate is a unitless percentage. This does not have the same statistical application as a normal  $R^2$ , but it does provide an idea of how the model performs, specifically of comparing the relative performance of the two levels of aggregation. The measure of fit estimates can be found in Table 2.4. The county level provides better estimates for soy and wheat, but the CRD level provides better estimates for WSG, which provides a motivation for utilizing the CRD level of aggregation.

## II.9 Constrained vs. Unconstrained Models

With the model aggregated and running at CRD levels, analysis turned to the validity of the results. As written, the model should run unconstrained to choose the coefficients which best fit the data. Indeed, an unconstrained model is desirable because it may be considered more statistically valid, as it more closely reflects the relationships present in the data. However, in applied econometric work, constraints are often imposed if the results do not fit economic theory or if the data may not perfectly reflect the situation faced by the decision maker, as in the case of proxies. To see if such constraints were necessary, the model was run unconstrained, and the elasticities (as calculated from (2.7) above) of the net returns variables were examined after imposing a .05% acreage

share threshold, because areas with small allocations to crops could skew results by showing very large elasticities for relatively small gross effects. Figure 2.3 is a histogram of the elasticities for corn net returns to corn acres planted. While the majority of CRDs and years follow theory, with a low, positive impact, there is a not-insignificant portion of areas and years which have a negative impact, which does not follow theory. This is seen further in examining soy net returns to corn acres (Figure 2.6): there is a much wider spread, and the majority of elasticities are positive. While this may suggest strong soy/corn cropping patterns, it is difficult to justify the assertion that soy acres will increase when the net returns to corn increase.

WSG elasticities are likewise problematic. While the relationship between WSG and crops such as corn and soybeans may be argued, there should be little doubt that it should have a positive own net return elasticity. Figure 2.7 shows the elasticities of pastureland rental rates to WSG area. The spread is extremely broad, and over 20% of the elasticities are strongly negative. This may indicate a proxy problem, but, given the location of WSG, pastureland rental rates were determined to be the best available proxy, as suggested earlier. Pasture rental rates represent a poor proxy for WSG returns in an area if it is the case that pasture and warm season grasses act as substitutes there, with uncultivated WSG replaced by seeded pasture in the event that pasture returns rise. These elasticities might also indicate other issues, such as autocorrelated error terms, for example from a multi-year drought. The exact cause is unknown.

In light of the results from Figures 2.3, 2.6, and 2.7, a series of constraints were imposed on the values of the coefficients estimated in (2.4). The elasticities were indirectly constrained by way of the coefficients themselves because elasticities are not

explicitly represented in this estimation method, but are calculated ex post instead. These constraints were designed to direct the model as little as possible while still ensuring the correct signs and relative magnitudes on own net returns elasticities. Given the known competitive nature of corn and soybeans, negative cross effects constraints were imposed on corn and soy net returns to soy and corn acres, respectively. Likewise, because of known soy-wheat double cropping patterns in a small part of the study area, cross effect restrictions are not applied between those crops. The constraints are not listed here: their indirect nature makes analysis uninformative.

For the rest of this paper, results presented will reflect the constrained model. The constrained net return elasticities can be seen in Figure 2.8. The elasticities fit theory much more closely, with positive own net return effects and mostly negative cross effects. The notable exception is soy/wheat, with mixed cross effects alluding to double-cropping. The elasticities indicate a much higher amount of double cropping than is known to exist in the area of study. However, the amount by which one should constrain the cross effects is difficult to ascertain. These elasticities are admittedly not completely ideal. For instance, the corn net return response to WSG is too strongly negative, greater than the own return response. The greater magnitudes of cross-effect elasticities can present problems with multiple simultaneous shocks. That is, if both returns to corn and WSG increase by the same amount, it will cause a net decrease in both corn and WSG area due to the stronger cross effects. However, shocking one price or return at a time should perform following theory, with decreases in cross purposes and increases in own purposes to positive shocks. With the correct directional impacts following a shock to one expected net return variable, area allocation will also move in the correct direction given

exogenous total area. There remain several potential directions for improving the constraints imposed here, such as by targeting specific elasticities suggested by previous literature on corn, soybean, and wheat area response or else by ensuring more theoretical constraints, such as symmetry. However, the directional impacts are mostly correct and the focus of this work on WSGs undermines somewhat the value of basing parameters narrowly on previous studies of the main crops.

It should be noted that forcing the model to reflect expectations did not come without cost. While the economics are more closely followed, the constraints impact the statistical validity of the errors of the coefficients derived from the covariance matrix in (2.8). For this reason, confidence intervals and t-tests are not provided for the coefficients. Further, the ability of the model accurately to estimate acreage is negatively impacted. Table 2.5 shows the difference in gross errors (2.9) between the constrained and unconstrained models. While the constrained model generally does poorer, it is worse with the estimation of WSG: the difference in WSG errors is more than three times greater than the next highest difference in errors. While this is unfortunate, the constraints are nonetheless imposed to achieve results which, when shocked with price or policy changes, will more closely follow economic theory in response, if not in actual acreage. Because of the loss of goodness of fit, the rest of the paper will focus on relative changes in acreages based on shocks to the model.

## **II.10 WSG Returns Shock and Area Response**

With the model area and constraints defined, the focus of this paper changes to WSG responses. To provide a frame of reference, Figure 2.9 shows the historical share of



land allocated to each CRD, averaged over time. To show how WSG responds to changes, the model was simulated dynamically over the sample period, calculated iteratively using (2.5). To see how WSG responds to a significant shock in prices, PRR was doubled for all CRDs and years in the sample period, and the model was run again. Figures 2.10-2.13 show the differences in area between the baseline scenario and the increased PRR scenario, averaged over time. Darker WSG areas indicate greater increases, while corn, soy, and wheat area are darker as more land is lost to those uses. Each figure has a second figure with it (e.g. 2.10a) that shows the difference in area as a percent of the crop's base acres. This shows relative changes as opposed to gross changes.

The results are surprising in a couple ways, but before discussing results, it should be noted that specific acreage transitions cannot be assumed from these maps. Because of the probabilistic nature of the model, shifts in acreage come from most or all of the other sources at once, and these maps are representative purely of the net effects. While discussion of such individual transitions is possible, it would vary based on CRD and over time, and such in-depth analysis falls beyond the scope of this research. Here, specific transitions will only be discussed in the broadest of terms.

Doubling WSG returns while holding all other crop returns constant causes WSG to increase proportionally most in areas traditionally allocated to strict corn/soybean rotations: namely Iowa, southern Minnesota, and parts of Illinois, as shown in Figure 2.10. This result is important: it shows that WSG may indeed compete with traditional crops for land area if the prices for WSG are high enough, at least in this representation of land allocation. There were small relative increases in the western part of the sample

area, with little to no increases in WSG in the upper northern area covering upper Minnesota and Wisconsin. One odd result is the sharp border between northern Missouri and southern Iowa. Southern Iowa shows strong increases in WSG growth, increasing area allocated to WSG by 20-25% of total land, but immediately across the border, the increases call to between 4-7%. The pastureland rental rates are about \$8 higher on average in that area of Iowa than those in north Missouri, resulting in a \$16/acre higher rate under the shock scenario, which may explain the differences in the two allocations.

Wheat response is shown in Figure 2.11. Changes in wheat area are strongest in areas where wheat is traditionally grown: North Dakota, South Dakota, and western Kansas. Nebraska's response is not as strong, and the middle and eastern parts of the sample area show little to no land going leaving wheat production, with the notable exception being the area around the bootheel of Missouri and southern Illinois. But, even there, the response is modest. The relative response in Illinois is somewhat stronger, but that owes to the lower amount of acreage dedicated to wheat production.

Soybean area is shown in Figure 2.12. As mentioned prior, some areas historically allocated to corn and soybean rotation were planted with WSG, and those areas are also shown here, with Illinois, Iowa, and southern Minnesota losing soybean acres. However, the loss of soybean land is not restricted to corn/soy areas. The eastern Dakotas and western Minnesota also lose soy area. This may be indicative of changes in wheat/soy crop rotation patterns, given the modest wheat losses there as well.

Corn area is shown in Figure 2.13. The corn loss area is fairly concentrated in the middle and eastern parts of the sample area, in gross terms. In relative terms, most areas lose a significant portion of area allocated to corn.

For a broader look, changes in acreage devoted to each crop from the base scenario and shocks can be seen in Table 2.6. Comparing the first and second columns of data show the total regional impact of doubling pasture net returns on the allocation of land. Substantial land is taken away from corn and soybeans, but a smaller percentage from wheat. This suggests, contrary to prior expectations, that corn and soy compete more easily than wheat for a doubling of pasture returns. This could also be attributable to the relatively lower pastureland rental rates in the Great Plains states. The rates average between \$8 and \$20 over time in ND, SD, NE, and KS, while averaging \$20 to \$35 in IA, WI, IL, and most of MO. Doubling pastureland rental rates would therefore have a greater impact in those states, where corn and soy are historically grown. With this in mind, the impact on those areas are likely much overstated from the doubling of pastureland rental rates. However, this shock was not intended to provide historical context, merely to show that this model estimated with these data can generate acreage shifts, which it does. Indeed, given the magnitude of changes that did occur, even when areas with higher PRR are considered, the model shows that a much more modest change in PRR may cause land to be allocated from traditional crops. Further testing is needed to examine these effects more closely, and specifically the impact of shocks to other areas of the system. Thus, this research tests the model further by applying a more historically tenable shock.

## **II.11 Corn Net Returns Shocks and Area Response**

The initial shock scenario is intended to show how the model responds in regards to changes in WSG returns. The magnitude of the shock is admittedly strong, following

the earlier hypothesis that a strong change in returns to WSG would be necessary to shift land away from traditional crops such as corn/soy rotations, and is there to provide results which are clearly separable from other coterminous effects. It is useful, however, to see how the model responds to smaller, more likely changes in prices. Chen and Khanna (2012) estimate that by 2022, EISA mandates will cause a 23.5% increase in the price of corn, given a conservative estimate on the availability of sugarcane for ethanol production. To draw a connection between the price change of that study to the net returns used by the model of this study, the shock was multiplied by the average percent difference between prices and net returns in the historical period (93%). This method of assuming a proportional link between prices and net returns assumes implicitly that the prices of inputs represented in the corn cost data are also affected. This representation might be consistent with medium- or long-run relationships between corn output and input markets, not an immediate impact of a price change on net returns. To simulate changes in corn price, the model is run with a +/- 21.9% difference in corn net returns.

Figures 2.14-2.17 show the changes in area with respect to the lower corn net returns shock. It is clear that there are strong similarities between Figures 2.13 and 2.14. There is a difference in the scale of the two maps, likely a result of the change in the shock magnitude. But, the interesting similarity is that the areas of relatively strong change are the same across separate shocks. This remains the case for soy and wheat: with changes in scale, the strong areas of change for wheat are the Dakotas and Kansas, and the strong areas of change for soy are the eastern Dakotas, southern Minnesota, Iowa, Illinois, and northern Missouri. The higher corn net returns shock follows the same pattern. Therefore, it would be unhelpful to show the same maps again. WSG are

somewhat of an exception. The strong areas still show change, but there is expansion of lands in which WSG is grown, shown in Figure 2.15. More acres are allocated relative to other areas in Nebraska and Missouri than in the WSG returns shock, with less in central and northwestern Iowa. Figure 2.18 shows the changes in corn area from the higher corn net return shock. The higher corn net returns maps look the same as the lower (Figure 2.14), with a change from gains in area to losses. This pattern follows for all other higher corn net return maps. Therefore, they will not be presented.

The important difference between the maps are the legends, showing the difference in scales. Since it has been established that, for the most part, the relative strength of changes by crop shown in Figures 2.14-2.17 are constant across scenarios, it may be easier to examine the impact of the shocks by examining the data by other means. A histogram of the shock share differences by crop and scenario, averaged over time, are given in Figure 2.19. The shocks, for the most part, follow economic theory in that increases to corn net returns increase corn area while decreasing other areas. Wheat does look odd, with what appears to be a not-insignificant portion of the area on the wrong side of each shock scenario. However, each of these are small enough ( $<.001$  in absolute value) that they may be considered effectively zero. The means of the differences bear this out. The histogram shows also that WSG area does react to changes elsewhere in the economy. This further supports the possibility that WSG could compete with traditional crops, based solely on changes in the other crops' own net returns.

For a more numerical analysis, Table 2.6 also shows the regional results of the two corn net return shocks. Area moves in the expected direction at the regional level, with higher corn net returns causing increases in corn area, and decreases to other crops.

The opposite occurs when corn net returns decrease. The model estimates have WSG as very elastic overall: a 20% change in corn net returns results in 40% changes in WSG area. This indicates that further refinement of the constraints may be needed, or it could point to problems with the use of pasture rental rates as a proxy for WSG net returns at the margin. Also, it may prove interesting to run this model at the national level, as opposed to the CRD level, to apply more relevant national policy shocks.

## **II.12 Conclusion**

This research sought to examine the relationships between WSG and traditional crops, namely corn, soybeans, and wheat. To accomplish this, WSG data were obtained from NASA satellite images and combined with NASS, ERS, and grain elevator data for corn, soybeans, and wheat. These data were placed in a dynamic MDP to estimate changes in land use, given a change in returns to WSG. However, issues arose with the estimation of coefficients and elasticities. The data did not fully cover the originally intended area, the computer could not solve a model combining CRD level coefficients and county level dummies, and initial results did not follow economic theory. To address these concerns, area of study was respecified, the level of share and explanatory data aggregation was changed to CRD, and constraints were imposed on the coefficients. The respecified model had a decreased goodness of fit, but it allowed examination of relative changes in acreage in response to a price shock.

Model performance was tested by changing key exogenous data. WSG and corn returns were changed over the historical period. WSG net returns were doubled to test model performance to a very large shock, in particular to see if this specification could permit exploration of the possibility that WSG competes with the main crops in the

region. In this shock, acres allocated to the different uses moved in the expected direction, with area being appropriately lost or gained from each of the major crops. This result shows that, if WSG is grown as a dedicated source for cellulosic biofuel, the model approach explored here could represent the degree to which it competes for land with other uses. However, the MDP, while initially thought to be useful in the estimation of limited-data scenarios, proved to be problematic in this case. In particular, the size of the impacts for the very large shock to the proxy for WSG net returns implied land use changes that seemed implausibly large, reducing regional corn, soybean, and wheat area dramatically. The corn net return shocks also provided directional effects that are correct, but no comparison is made with other studies to confirm that the magnitudes conform to findings found in other studies, and the impacts of the corn shocks on WSG area suggest that the model might over-state the substitution of WSG and other land uses, or at least corn use. Further research could take this data sample and apply it with more traditional models to compare results. Alternatively, refinements in the estimation method to take into account spatial interdependence or impose parameter restrictions that require elasticities to fall with ranges indicated by other research, where relevant, could be undertaken.

## II.13 References

- Adelman, I.G. 1958. "A Stochastic Analysis of the Size Distribution of Firms." *Journal of the American Statistical Association* 53: pp. 893- 904.
- Ahn, S., A.J. Plantinga, and R.J. Alig. 2000. "Predicting Future Forestland Area: A Comparison of Econometric Approaches." *Forest Science* 46(3): pp. 363-376.
- Brechbill, S.C., W.E. Tyner, and K.E. Ileleji. 2011. "The Economics of Biomass Collection and Transportation and Its Supply to Indiana Cellulosic and Electric Utility Facilities." *Bioenergy Research*. 4: pp. 141-152.
- Burnham, B.O. 1973. "Markov Intertemporal Land Use Simulation Model." *Southern Journal of Agricultural Economics* 5(1): pp. 253-258.
- Carpenter, B. and M. Brees. 2010. "Switchgrass for Biofuel Projected Budgets for Crop Established 2011." University of Missouri Dept. Agr. App. Econ., October.
- Chen, X. and M. Khanna. 2012. "Food vs. Fuel: the Effect of Biofuel Policies." *American Journal of Agricultural Economics* 95(2): pp. 289-295.
- De La Torre Ugarte, D.G., M.E. Walsh, H. Shapouri, and S.P. Slinsky. 2003. "The Economic Impacts of Bioenergy Crop Production on U.S. Agriculture." Washington DC: U.S. Department of Agriculture, OEPNU Ag. Econ. Rep. 816, February.
- Devadoss, S. and J. Bayham. 2010. "Contributions of U.S. Crop Subsidies to Biofuel and Related Markets." *Journal of Agricultural and Applied Economics*, 42 (4): pp. 743–756.
- Dicks, M.R., J. Campiche, D. De La Torre Ugarte, C. Hellwinckel, H.L. Bryant, and J.W. Richardson. 2009. "Land Use Implications of Expanding Biofuel Demand." *Journal of Agricultural and Applied Economics*, 41 (2): pp. 435–453.
- Duffy, M. and V.Y. Nanhou. 2001. "Costs of Producing Switchgrass for Biomass in Southern Iowa." Iowa State University Ext. PM 1866, April.
- Golan, A., G. Judge, and D. Miller. 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York NY: John Wiley and Sons.
- Golan, A., and S. Vogel. 2000. "Estimation of Non-Stationary Social Accounting Matrix Coefficients with Supply Side Information." *Economic Systems Research* 12(4): pp. 447-471.
- Hallberg, M. 1969. "Projecting the Size Distribution of Agricultural Firms: An Application of a Markov Process with Non-Stationary Transition Probabilities." *American Journal of Agricultural Economics* 51: pp. 289-302.



- Hardie, I.W., and P.J. Parks. 1997. "Land Use with Heterogeneous Land Quality: An Application of an Area Base Model." *American Journal of Agricultural Economics* 79(2): pp. 299-310.
- Heady, E. O. 1954. "Simplified Presentation and Logical Aspects of Linear Programming Technique." *Journal of Farm Economics* 36(5): pp. 1035- 1048.
- Helsel, Z.R., and J. Alvarez. 2012. "Economic Potential of Switchgrass as a Biofuel Crop in Florida." Dept. Farm Res. Econ. FE900, University of Florida.
- Hofstrand, D. and W. Edwards. 2003. "Crop Revenue Insurance." Dept. Econ. FM-1853, Iowa State University.
- Holman, J., R. Gillen, J.L. Moyer, T. Roberts, K. Harmoney, P. Sloderbeck, T. Dumler, K. Martin, S. Staggenborg, W. Fick, S. Maxwell, and C. Thompson. 2011. "Kansas Switchgrass Production Handbook." Kansas State University Agr. Exp. Res. Sta. MF3018, November.
- Jaynes, E.T. 1957. "Information Theory and Statistical Mechanics." *Physical Review* 106(4): pp. 620-630.
- Kelley, A.C. and L.W. Weiss. 1969. "Markov Processes and Economic Analysis: The Case of Migration." *Econometrica* 37(2): pp. 280-297.
- Khanna, M., B. Dhungana, and J. Clifton-Brown. 2008. "Costs of Producing Miscanthus and Switchgrass for Bioenergy in Illinois." *Biomass and Bioenergy* 32: pp. 482-493.
- Khanna, M. A. Jain, and A. Oliver. 2011. "Production of Bioenergy Crops in the Midwest." University of Illinois Urbana-Champaign Dept. Agr. & Cons. Econ. FEFO 11-06, April.
- Kim, C.S., G. Schaible, and S. Daberkow. 2010. "The Relative Impacts of U.S. Bio-Fuel Policies on Fuel-Energy Markets: A Comparative Static Analysis." *Journal of Agricultural and Applied Economics*, 42 (1): pp. 121–132.
- Lubowski, R.N. 2002. "Determinants of Land-Use Transitions in the United States: Econometric Analysis of Changes Among the Major Land-Use Categories." PhD Dissertation, Harvard University.
- Lubowski, R.N., A.J. Plantinga, and R.N. Stavins. 2003. "Determinants of Land Use Change in the United States 1982-1997." Discussion Paper 03-47, Washington DC: Resources for the Future, August.

- . 2008. "What Drives Land-Use Change in the United States? A National Analysis of Landowner Decisions." *Land Economics* 84 (4): pp. 529–550.
- McFadden, D., and K. Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15: pp. 447-470.
- McRae, E. 1977. "Estimation of Time-Varying Markov Processes with Aggregate Data." *Econometrica* 45: pp. 183-98.
- Miller, D.J. and A.J. Plantinga. 1999. "Modeling Land Use Decisions with Aggregate Data." *American Journal of Agricultural Economics* 81(1): pp. 180-194.
- Miller, J.C. and K.H. Coble. 2011. "Incentives Matter: Assessing Biofuel Policies in the South." *Journal of Agricultural and Applied Economics*, 43 (3): pp.413–421.
- Paap, R. and H.K. van Dijk. 2003. "Bayes Estimates of Markov Trends in Possibly Cointegrated Series: An Application to U.S. Consumption and Income." *Journal of Business & Economic Statistics* 21: pp. 547-563.
- Perrin, R.K., K.P. Vogel, M.R. Schmer, and R.B. Mitchell. 2008. "Farm-Scale Production Cost of Switchgrass for Biomass." *Bioenergy Research* 1: pp.91-97.
- Public Law (PL) 110–140. 2007. Energy Independence and Security Act of 2007. [frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110\\_cong\\_public\\_laws&docid=f:publ140.110.pdf](http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ140.110.pdf). March 10, 2008.
- Ray, D.E., D.G. De La Torre Ugarte, M.R. Dicks, and K.H. Tiller. 1998. "The POLYSYS Modeling Framework: A Documentation." Staff Paper, Agricultural Policy Analysis Center, Dept. of Agricultural Economics and Rural Sociology, University of Tennessee.
- Shannon, C.E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27: pp. 379-423, 623-656.
- Smith, R. 2009. "Switchgrass for Biomass." Dept. Agr. Econ., University of Kentucky, November.
- Theil, H. 1969. "A Multinomial Extension of the Linear Logit Model." *International Economic Review* 10: pp. 251-259.
- Tompkin, J.R. 1958. "Response of the Farm Production Unit as a Whole to Prices Response of the Farm Production Unit as a Whole to Prices." *Journal of Farm Economics* 40(5): pp. 1115-1128.

US Department of Agriculture, National Agricultural Statistics Service (USDA/NASS). 2012. "Crop Production 2011 Summary." Available at [usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047](http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047).

University of Tennessee Extension. 2009. "Guideline Switchgrass Establishment And Annual Production Budgets Over Three Year Planning Horizon." University of Tennessee Dept. Agr. Econ. AE 10-02, July.

Wang, C., F. Fritschi, G. Stacey, and Z.W. Yang. 2011. "Phenology-Based Assessment of Perennial Energy Crops in North American Tallgrass Prairie." *Annals of the Association of American Geographers* 101(4): pp. 742-751.

Wu, J., and K. Segerson. 1995. "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." *American Journal of Agricultural Economics* 77(4): pp. 1033-1047.

Wullschleger, S.D., E.B. Davis, M.E. Borsuk, C.A. Gunderson, and L.R. Lynd. 2010. "Biomass Production in Switchgrass across the United States: Database Description and Determinants of Yield." *Agronomy Journal* 102(4): pp. 1158-1168.

## II. 14 Appendix: Figures and Tables

Figure 2.1: Midwestern Land Covers Observed by MODIS, Indication of Presence for Six Out of Ten Years, 2000-2009

Source: Susan Wang, University of Missouri Geography Dept.

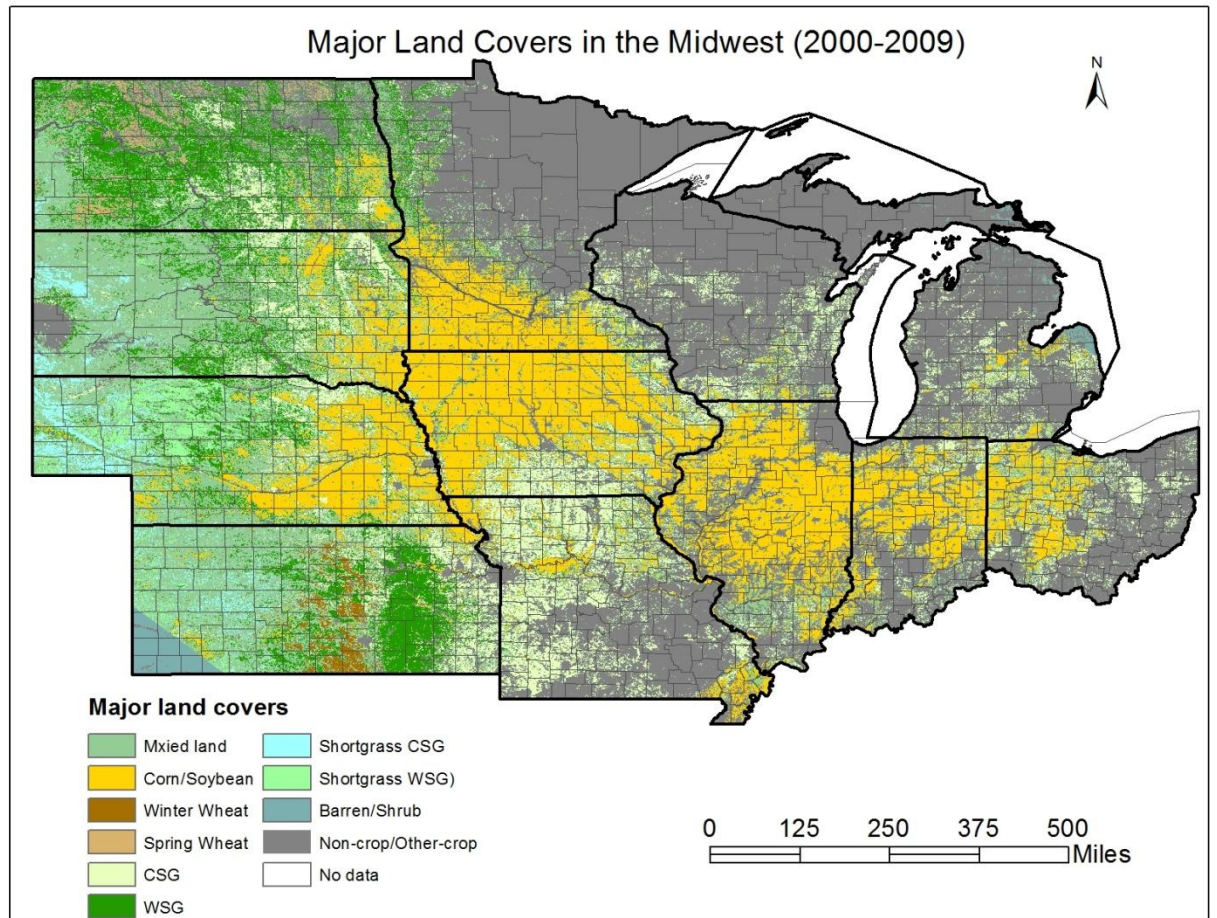


Figure 2.2: Corn Net Returns to Corn Area Planted Elasticities, County Level Aggregation, Unconstrained  
Source: Author's Estimates

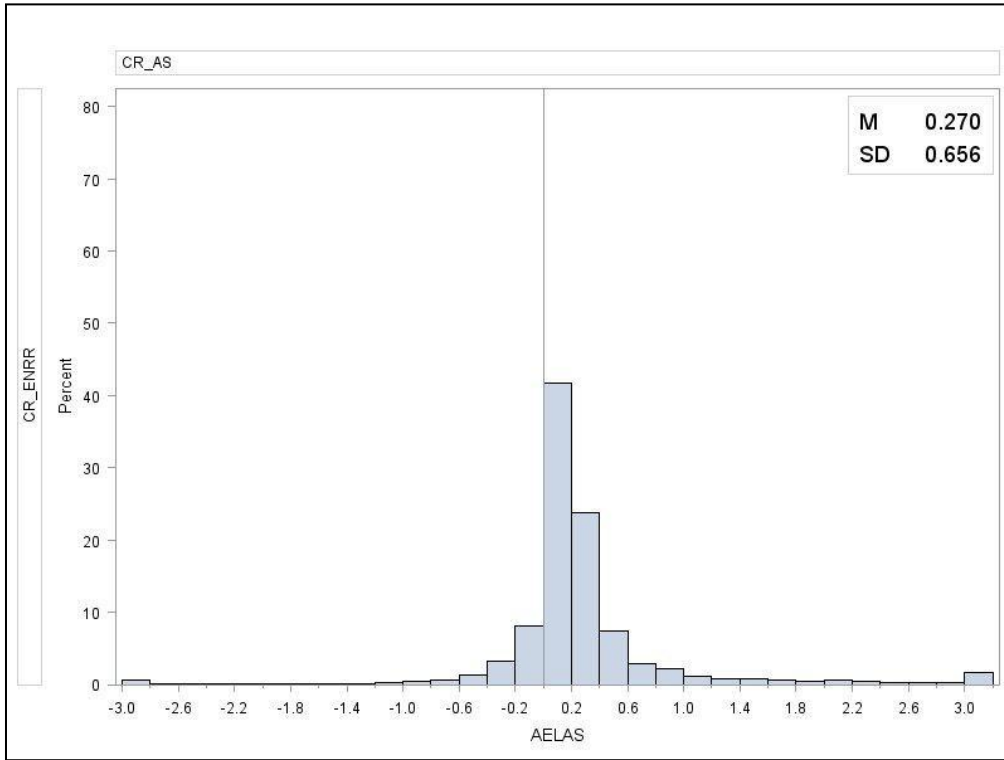


Figure 2.3: Corn Net Returns to Corn Area Planted Elasticities, CRD Level Aggregation, Unconstrained  
Source: Author's Estimates

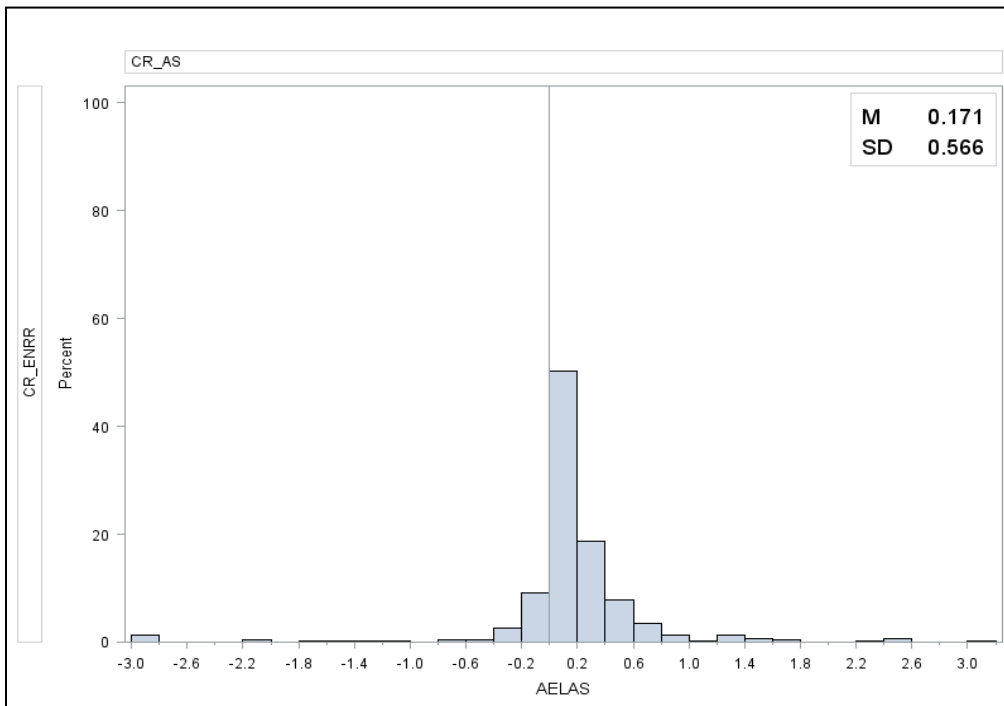


Figure 2.4: Pastureland Rental Rate to Wheat Area Planted Elasticities, County Level Aggregation, Unconstrained

Source: Author's Estimates

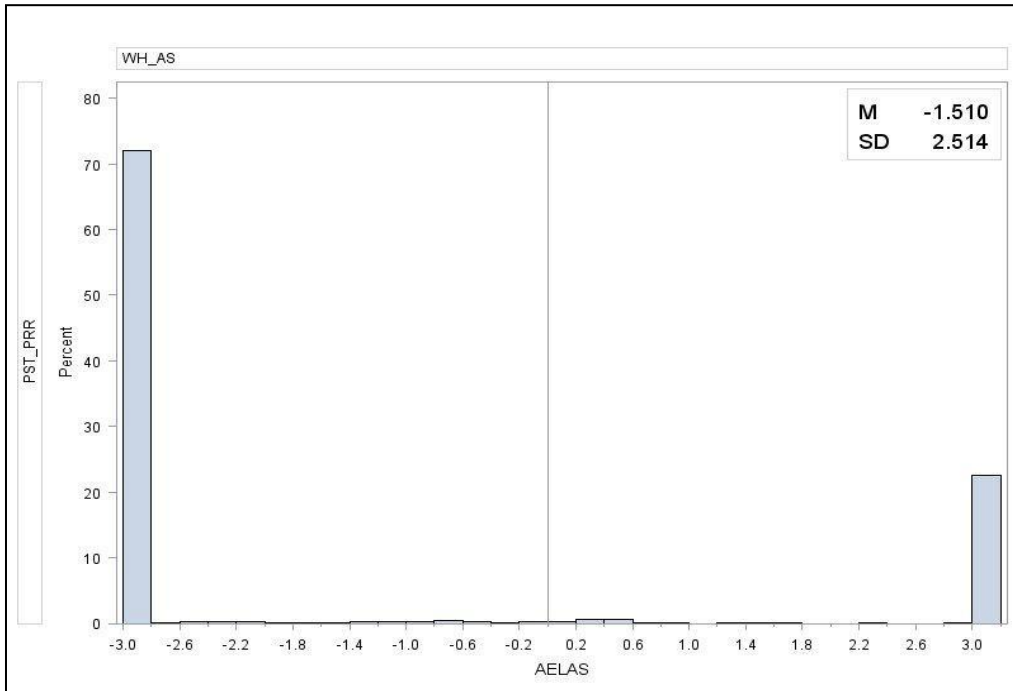


Figure 2.5: Pastureland Rental Rate to Wheat Area Planted Elasticities, CRD Level Aggregation, Unconstrained

Source: Author's Estimates

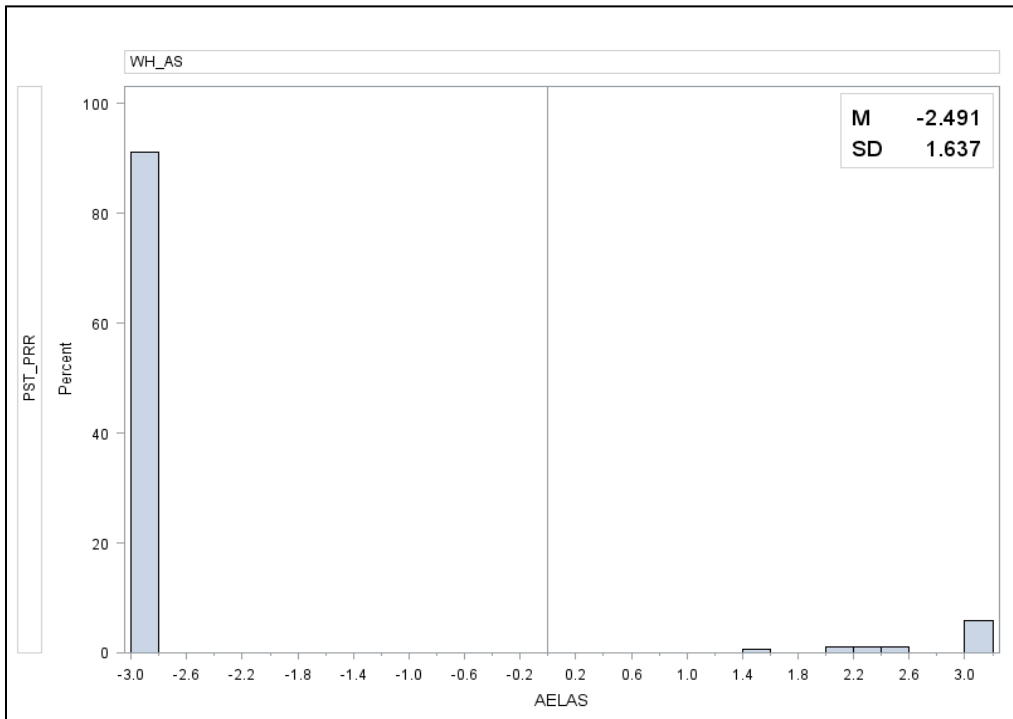


Figure 2.6: Soy Net Returns to Corn Acres Planted Elasticities, CRD Level, Unconstrained  
 Source: Author's Estimates

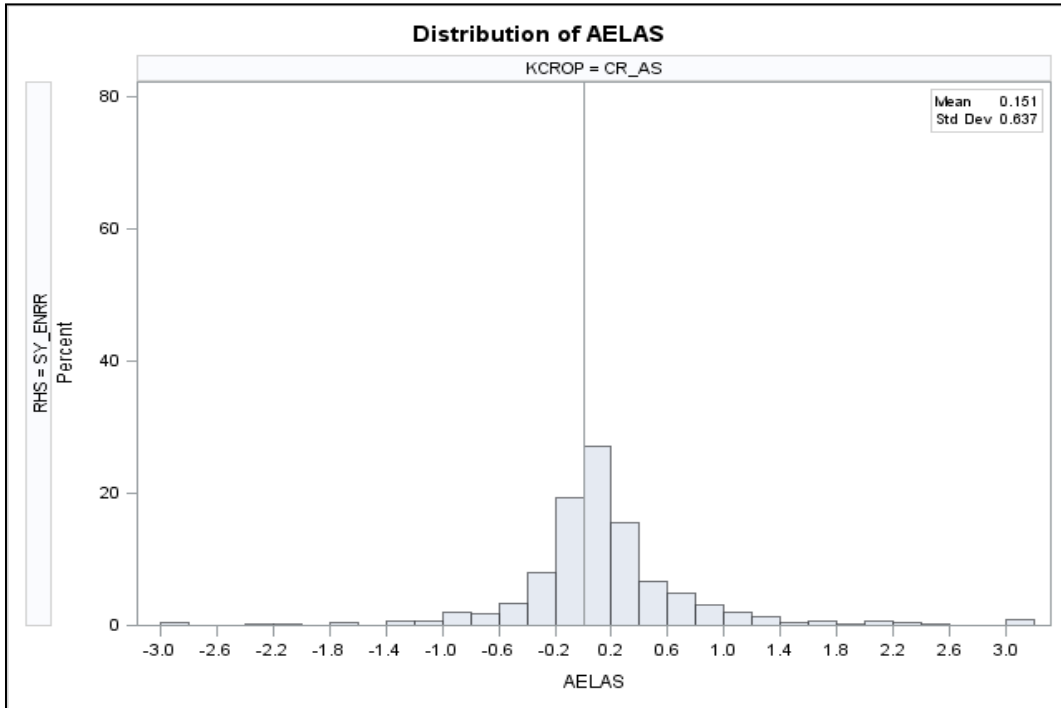


Figure 2.7: Pastureland Rental Rate to Warm Season Grass Acres Elasticities, CRD Level, Unconstrained  
 Source: Author's Estimates

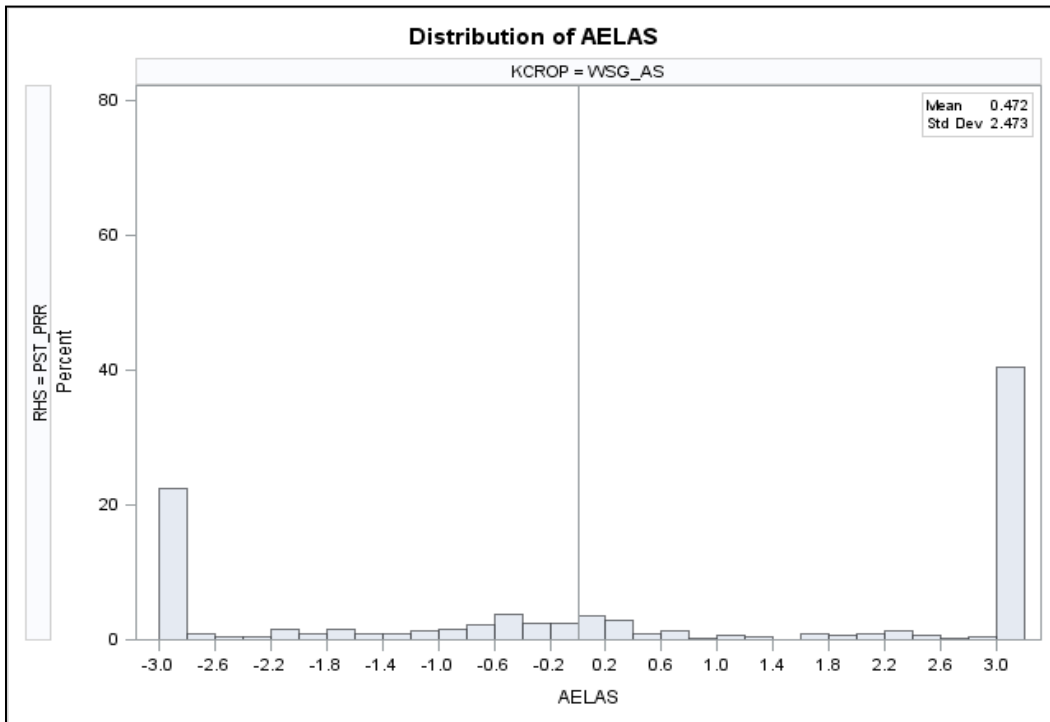


Figure 2.8: Net Returns Elasticities, CRD Level, Constrained Model  
 Source: Author's Estimates

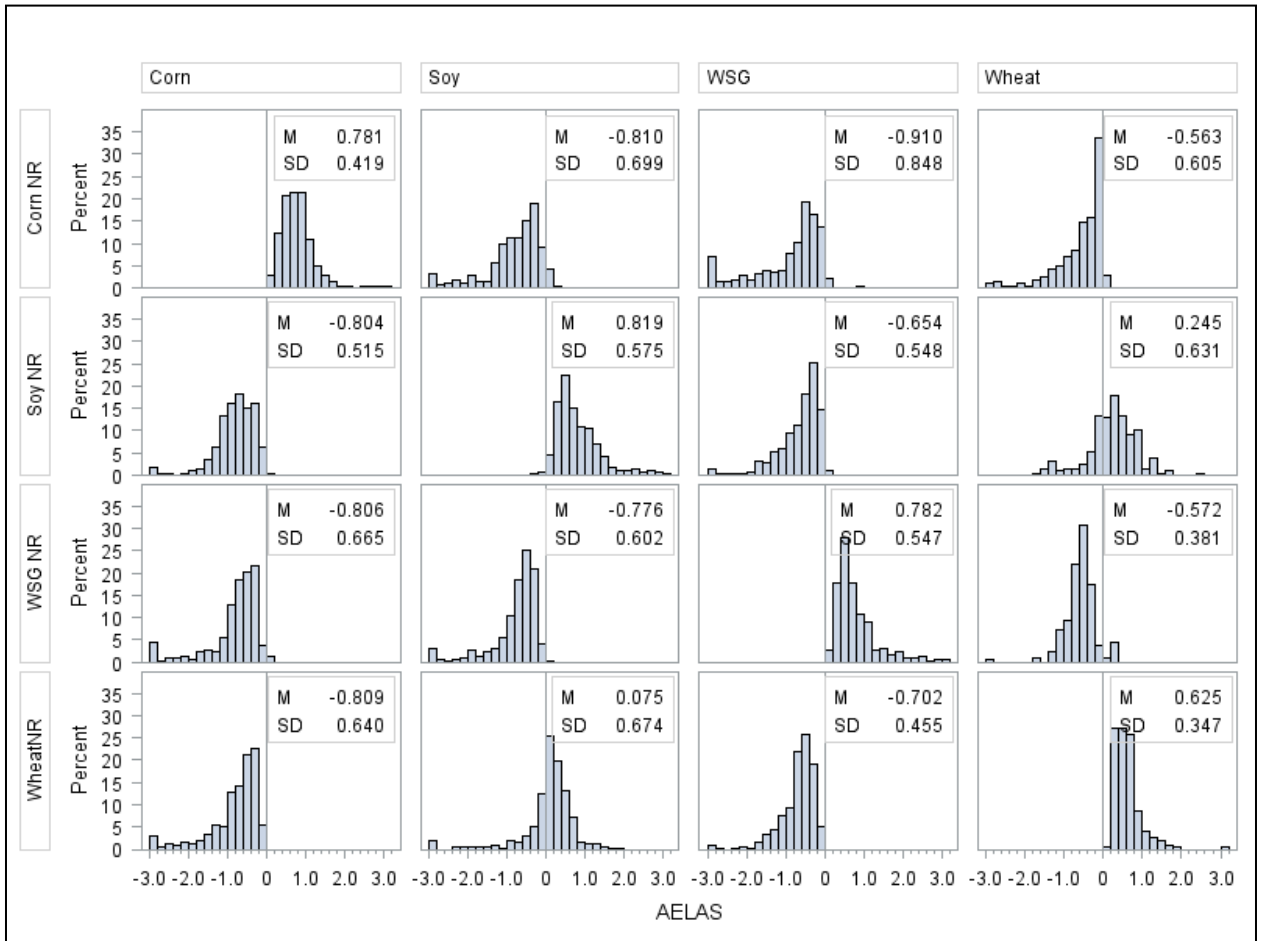




Figure 2.9: Historical Warm Season Grass Area by CRD as a Percent of Total Land, 2000-2009 Average  
Source: NASA MODIS Satellite Imaging Data

### WSG Area as a Percent of Total Land, 2000-2009 Average

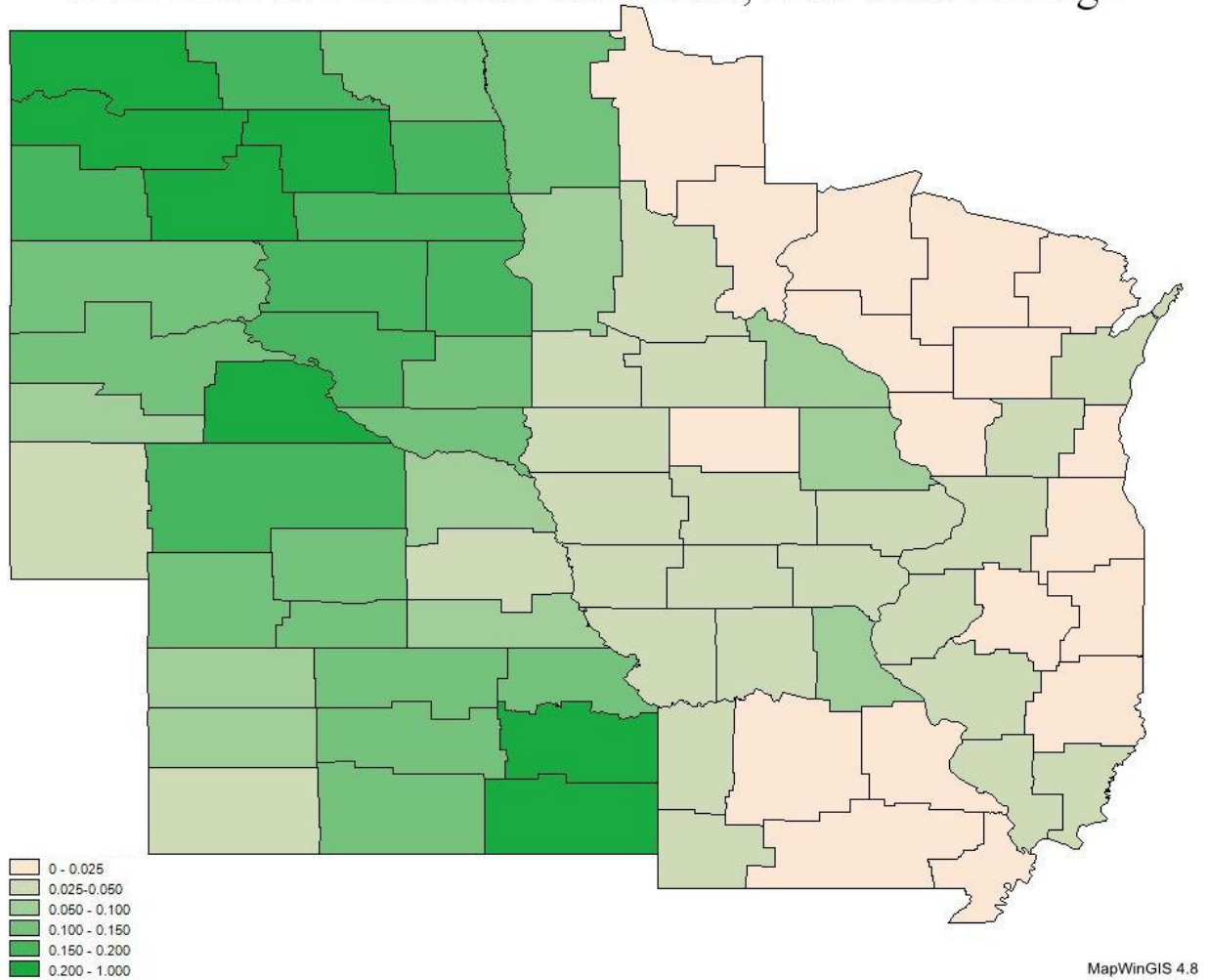


Figure 2.10: Change in WSG Area in Response to Doubled WSG Returns, 2001-2009  
 Average, Measured in 1000 Acres  
 Source: Author's Estimates

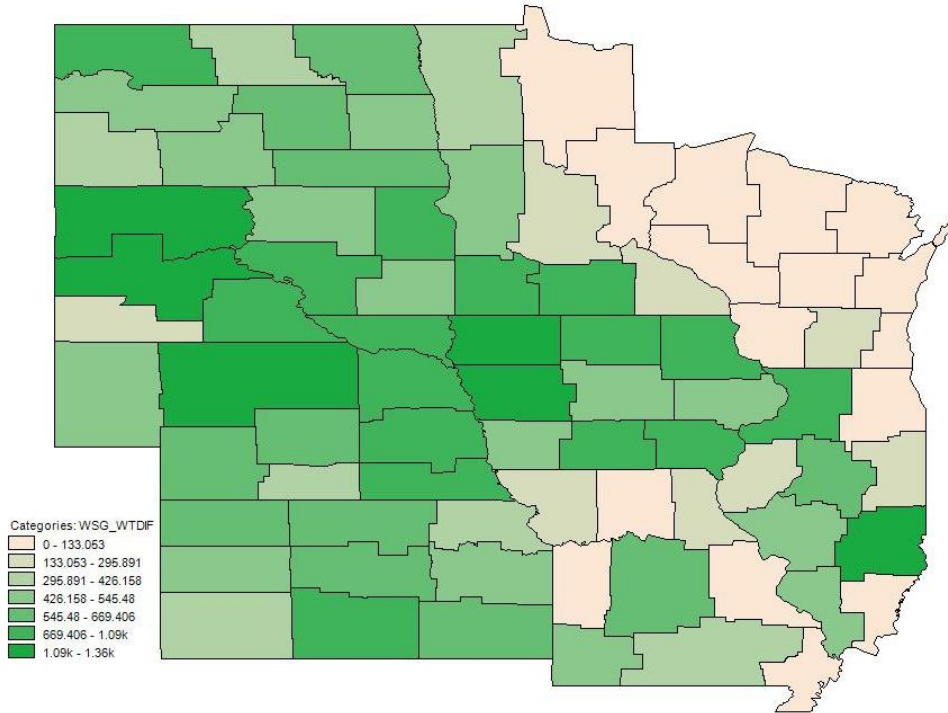


Figure 2.10a: Change in WSG Area in Response to Doubled WSG Returns, 2001-2009  
 Average, Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

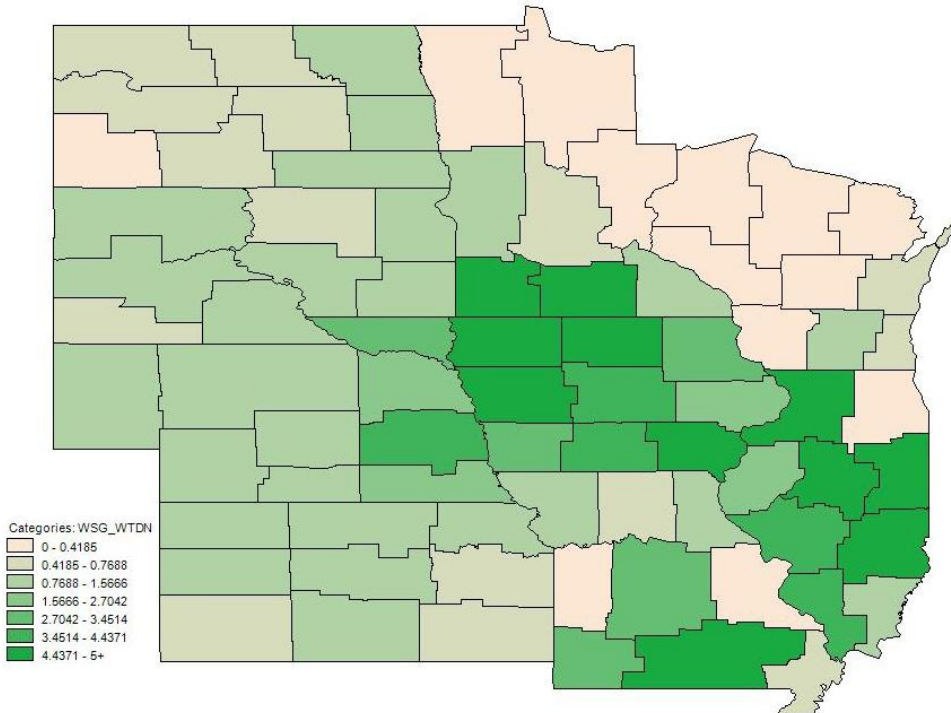


Figure 2.11: Change in Wheat Area in Response to Doubled WSG Returns , 2001-2009 Average, Measured in 1000 Acres  
 Source: Author's Estimates

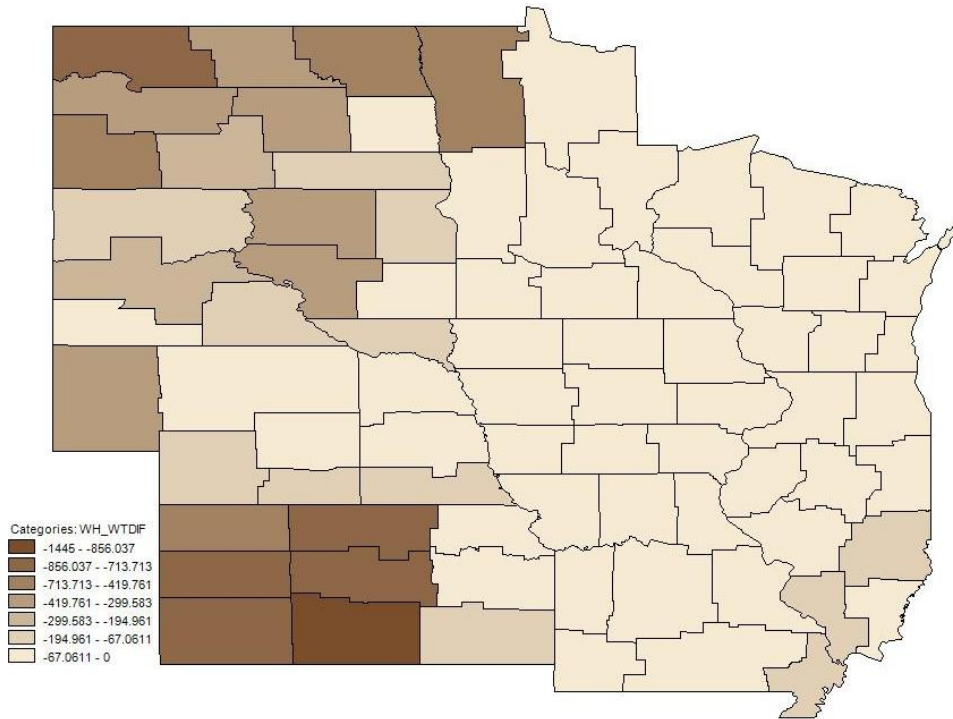


Figure 2.11a: Change in Wheat Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

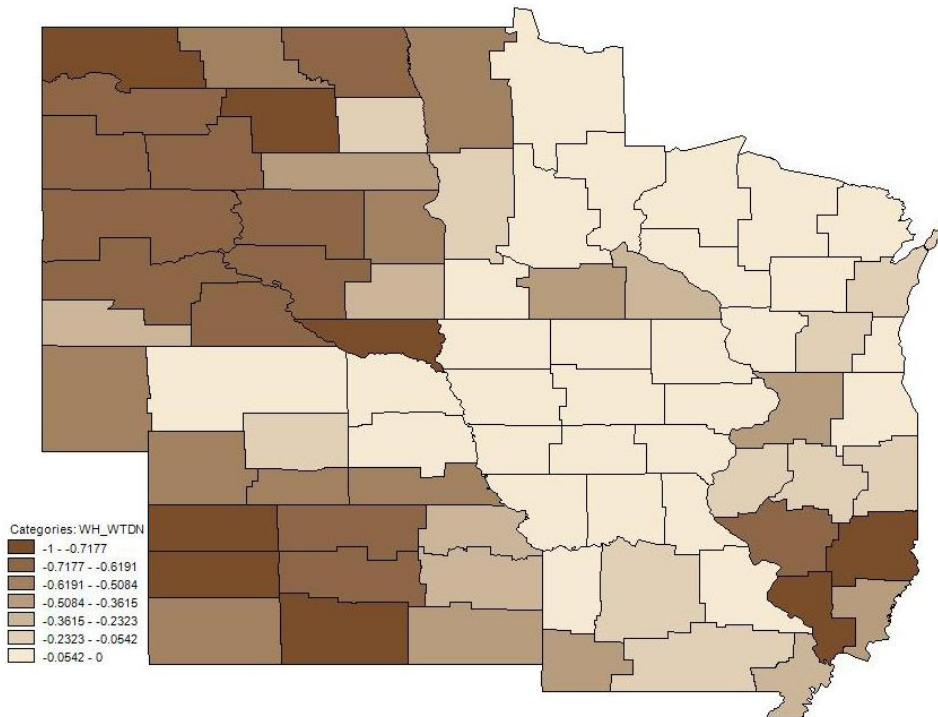


Figure 2.12: Change in Soybean Area in Response to Doubled WSG Returns, 2001-2009  
 Average, Measured in 1000 Acres  
 Source: Author's Estimates

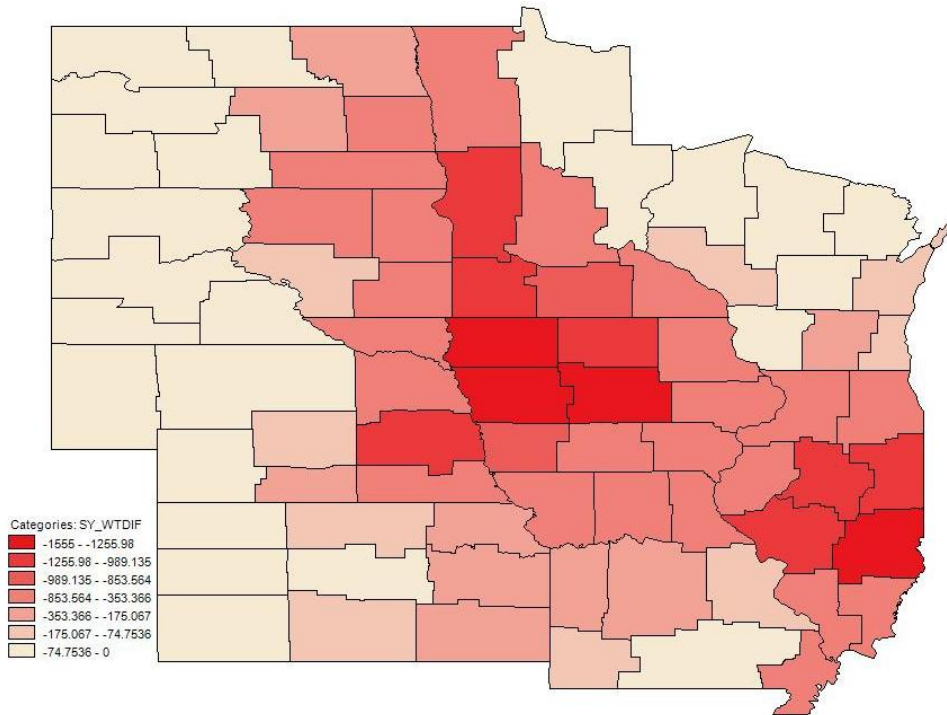


Figure 2.12a: Change in Soybean Area in Response to Doubled WSG Returns, 2001-2009  
 Average,  
 Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

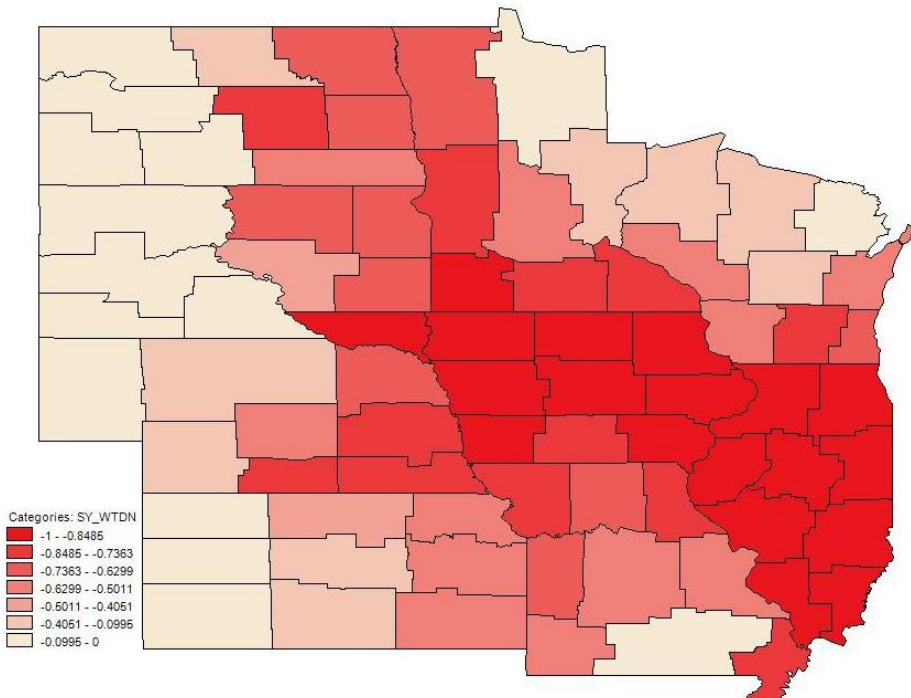


Figure 2.13: Change in Corn Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured in 1000 Acres  
 Source: Author's Estimates

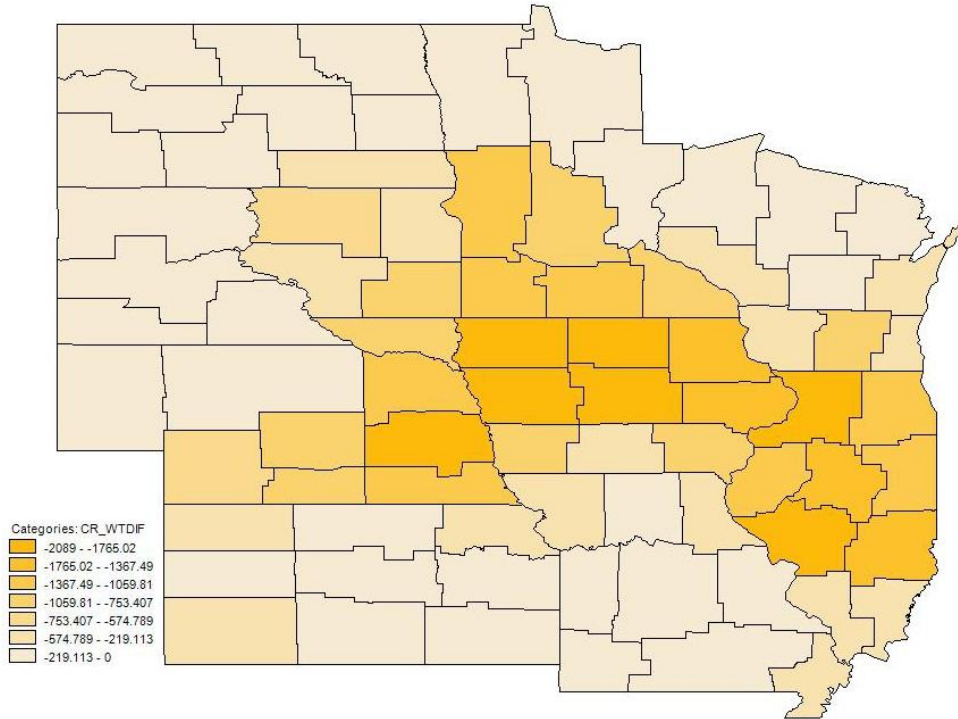


Figure 2.13a: Change in Corn Area in Response to Doubled WSG Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres

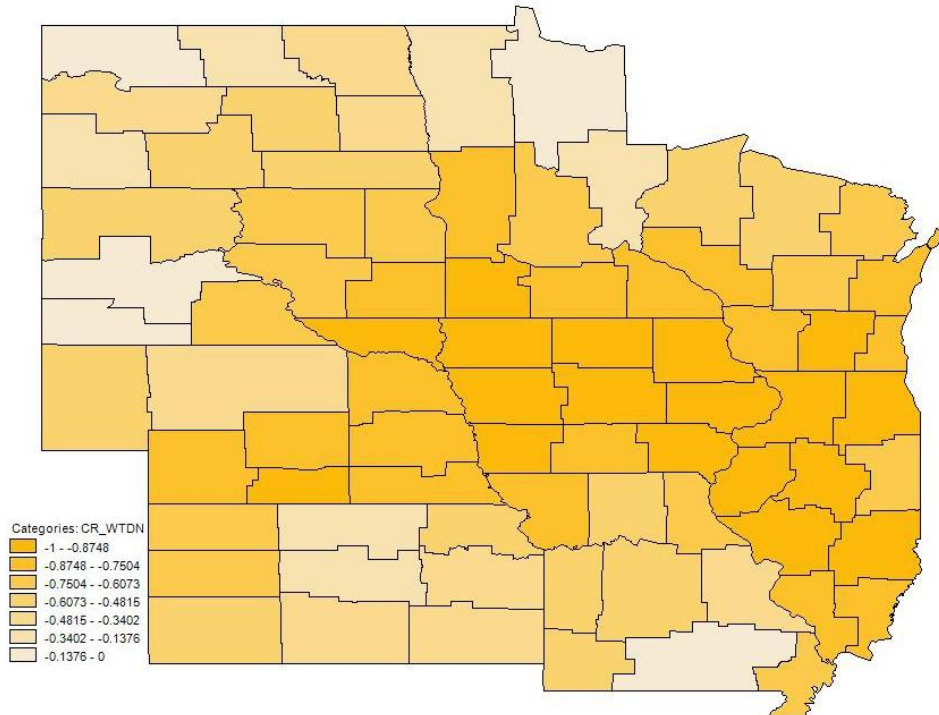


Figure 2.14: Change in Corn Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres  
 Source: Author's Estimates

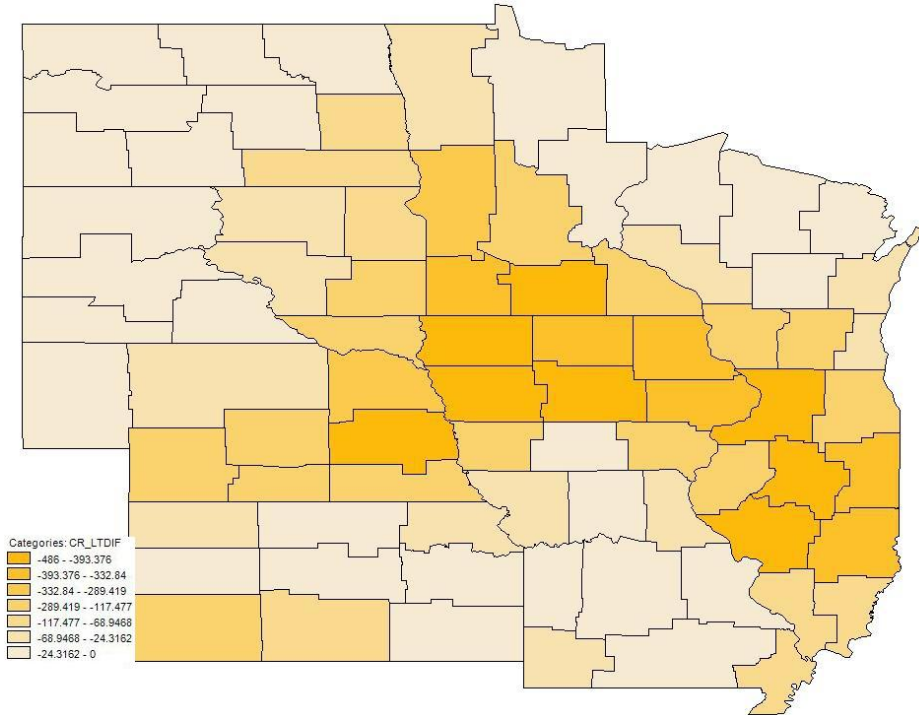


Figure 2.14a: Change in Corn Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

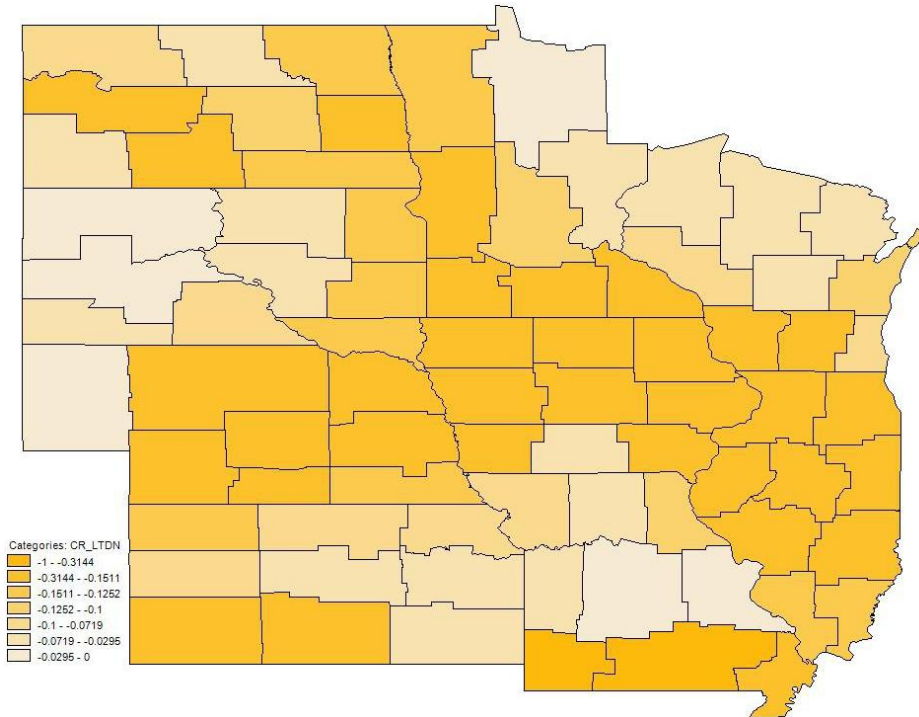


Figure 2.15: Change in Warm Season Grass Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres  
 Source: Author's Estimates

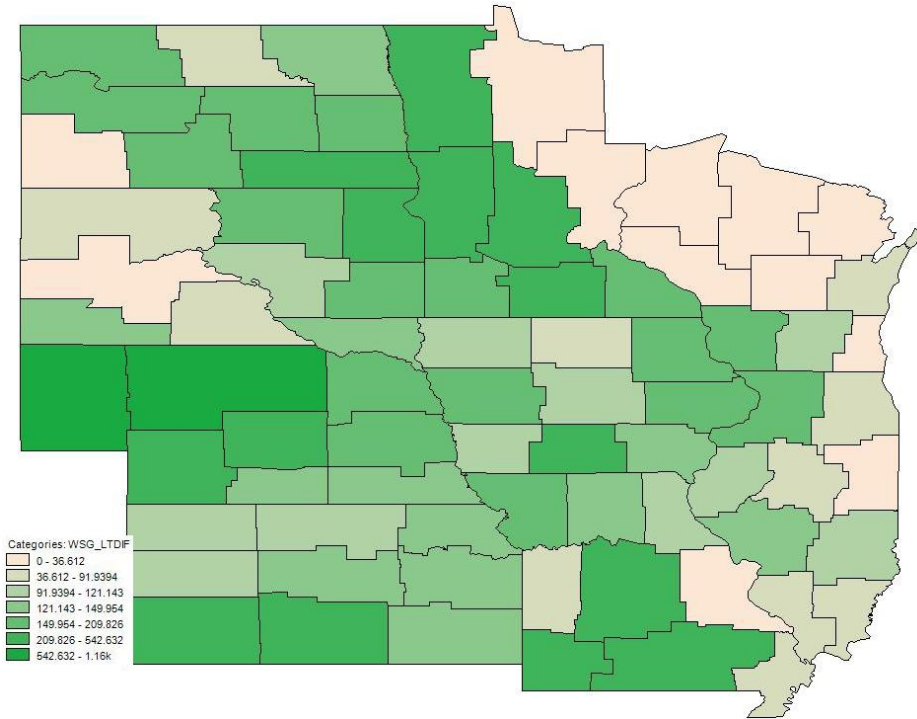


Figure 2.15a: Change in Warm Season Grass Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

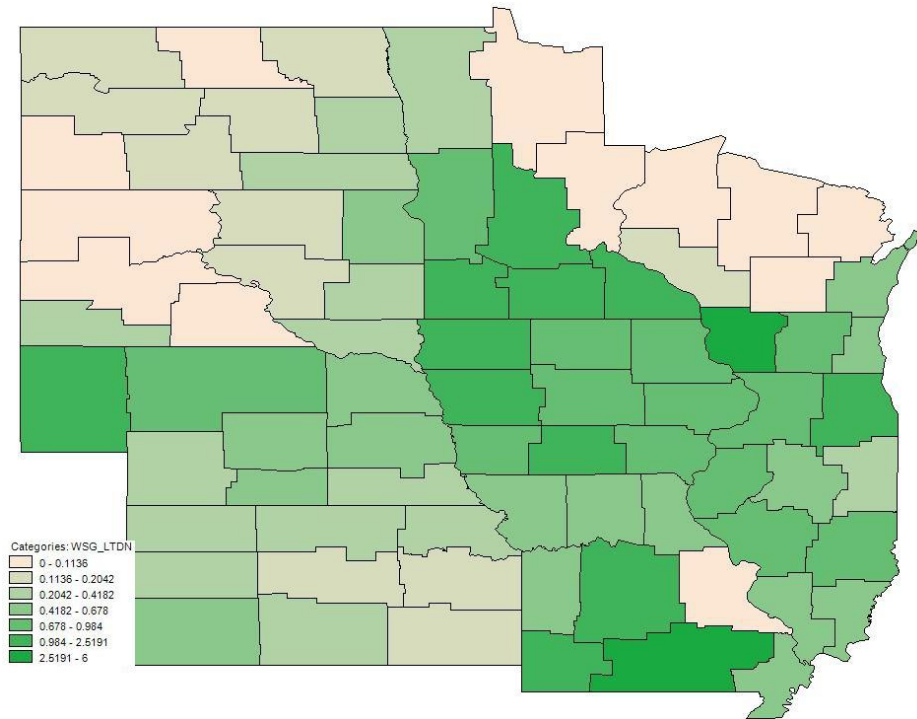


Figure 2.16: Change in Soy Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres  
 Source: Author's Estimates

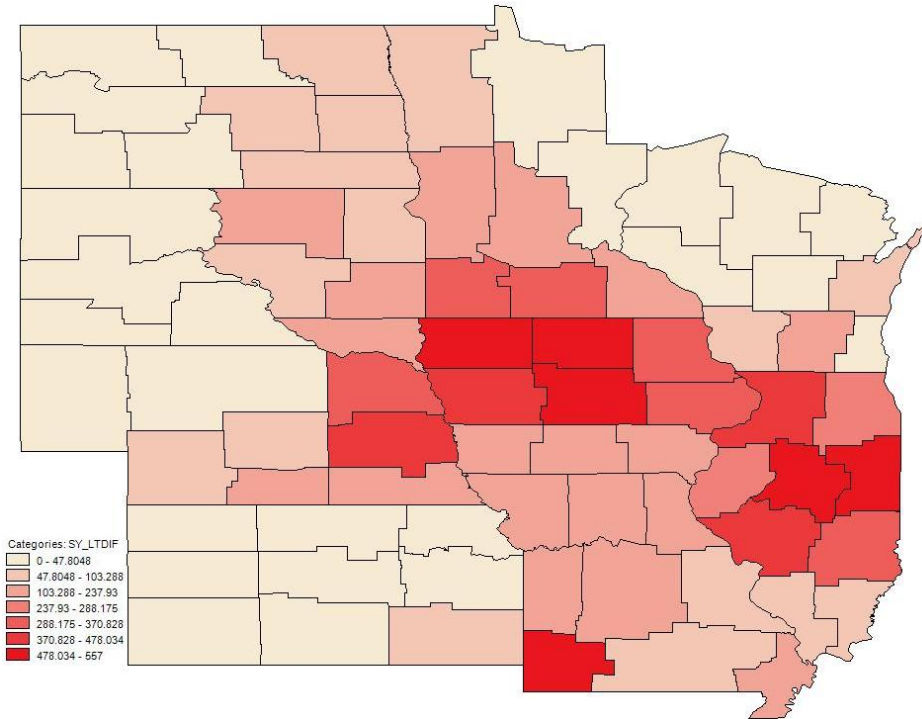


Figure 2.16a: Change in Soy Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

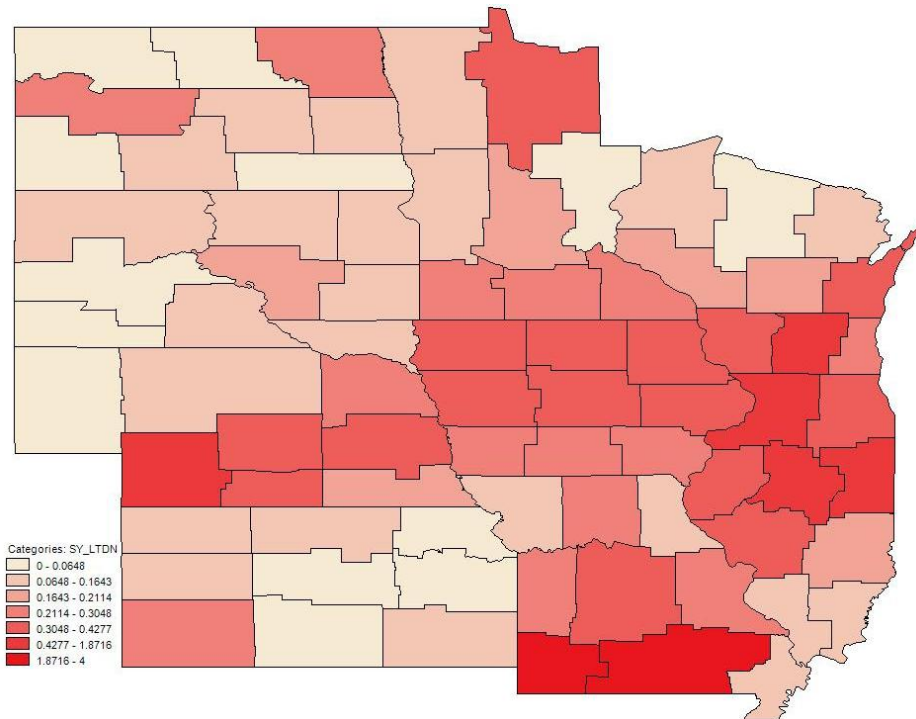




Figure 2.17: Change in Wheat Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres  
 Source: Author's Estimates

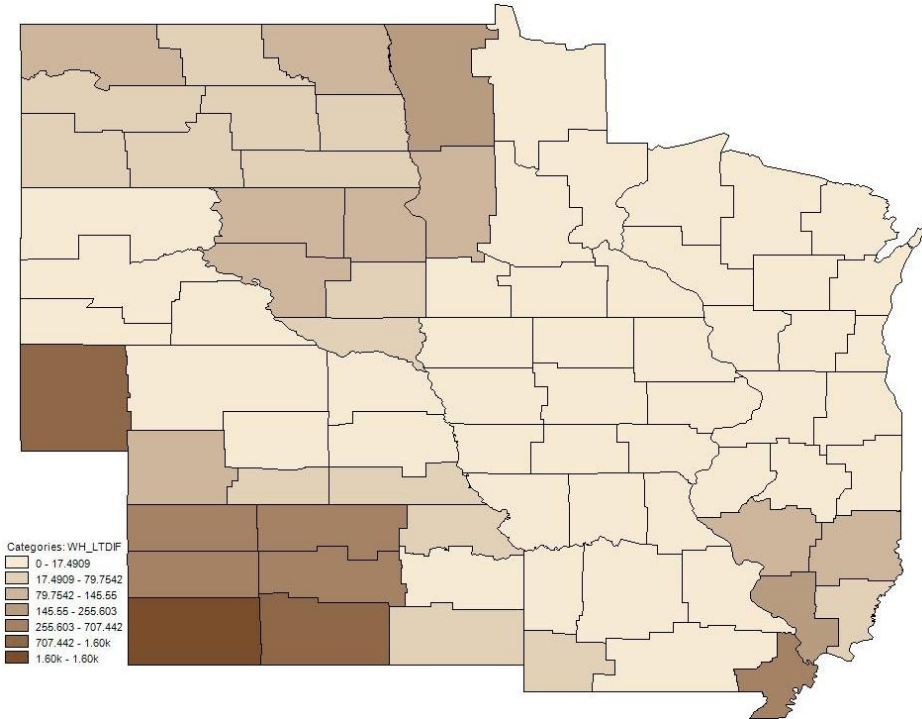


Figure 2.17a: Change in Wheat Area in Response to Lower Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

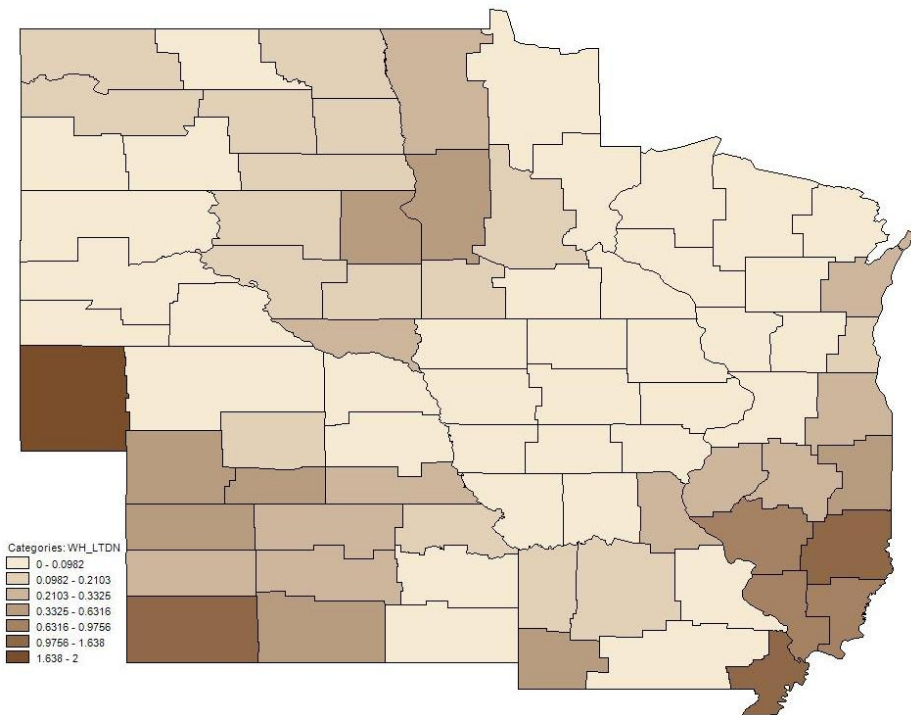


Figure 2.18: Change in Corn Area in Response to Higher Corn Net Returns, 2001-2009 Average, Measured in 1000 Acres  
 Source: Author's Estimates

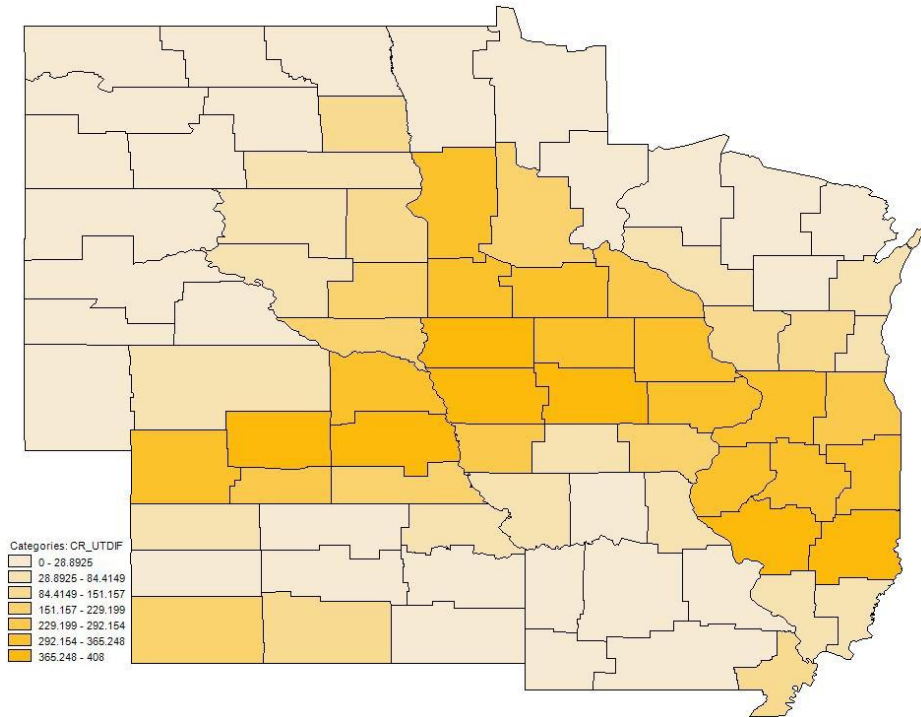


Figure 2.18a: Change in Corn Area in Response to Higher Corn Net Returns, 2001-2009 Average, Measured as a Percent of Base Scenario Acres  
 Source: Author's Estimates

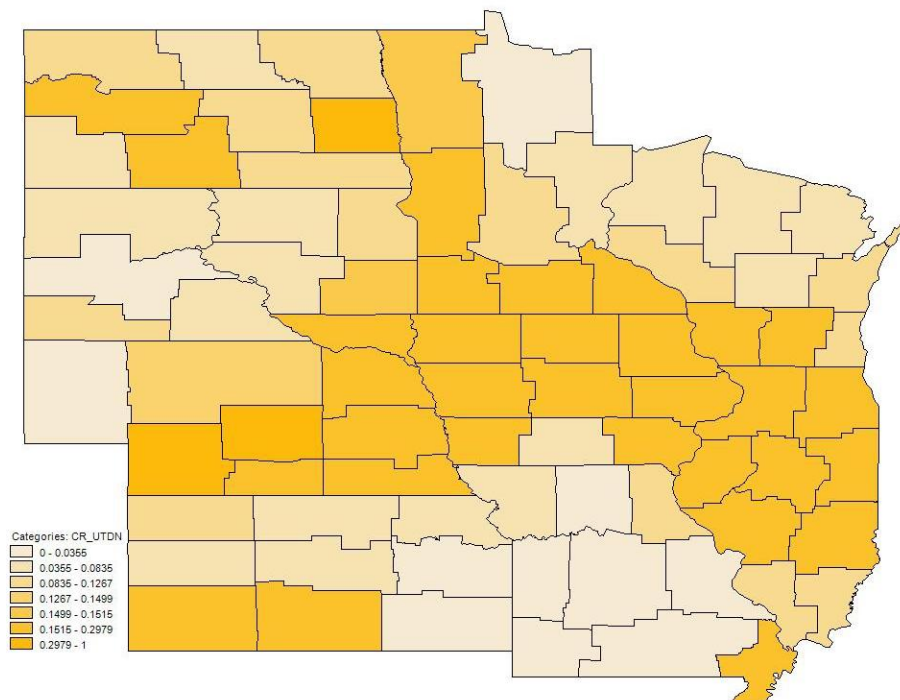


Figure 2.19: Histograms of Changes in Area Allocated to Each Crop as a Percent of Total Area by Scenario, +/- 21.9% in Corn Net Returns, 2001-2009 Average  
 Source: Author's Estimates

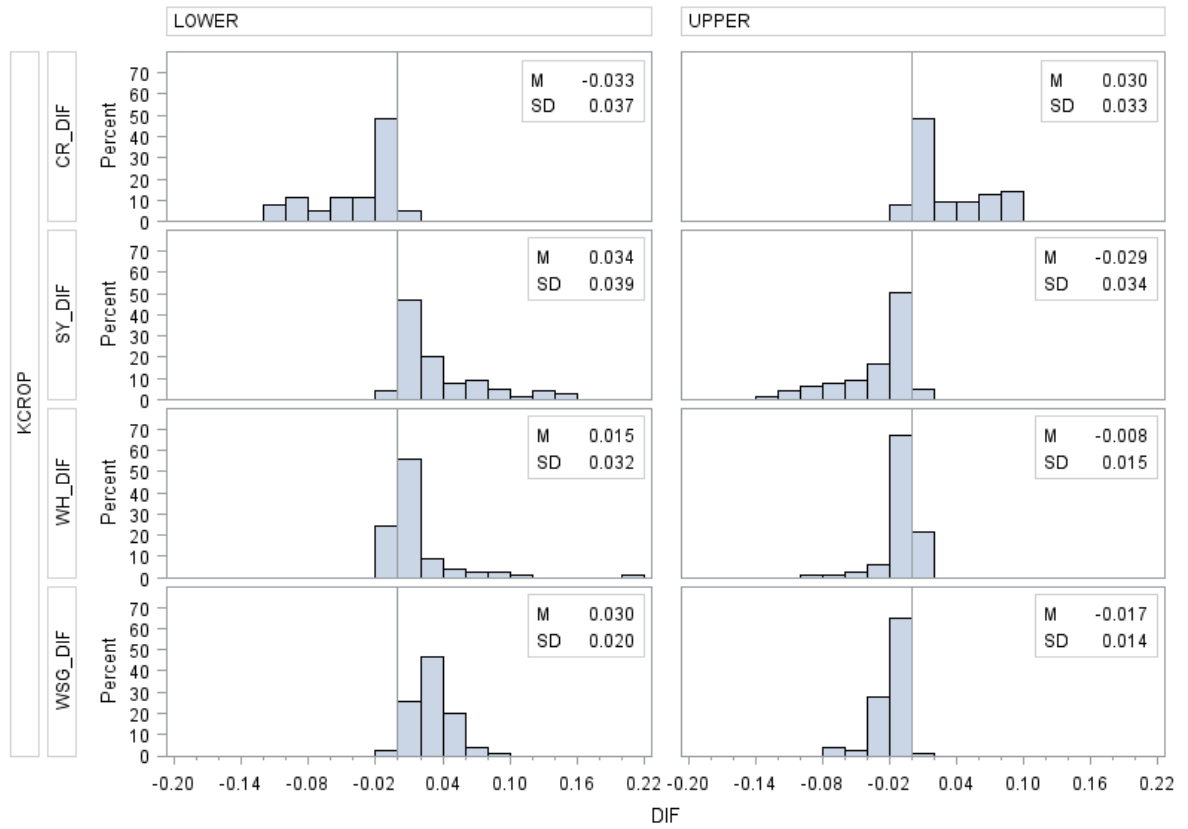


Table 2.1: Example Markovian Land Transition Probabilities, 2001 – 2002  
Source: Author's Estimates

2001	2002					Sum
	Corn	Soy	Wheat	WSG	Residual	
Corn	0.26	0.25	0.03	0.06	0.4	1
Soy	0.26	0.25	0.03	0.06	0.4	1
Wheat	0.22	0.22	0.16	0.17	0.23	1
WSG	0.22	0.22	0.16	0.17	0.24	1
Residual	0.23	0.2	0	0.01	0.56	1

Table 2.2: Shares and Expected Net Returns by Crop, Averaged Across Selected CRDs and All Sample Years

Source: Author's Calculations and Estimates Based on Observations Used for Estimation

Crop	Average Share (% of Total Area)	Average Expected Net Returns (USD/acre)
Corn	0.166	183.207
Soy	0.143	156.272
Wheat	0.052	91.405
WSG	0.075	21.849
Residual	0.565	

Table 2.3: Prior Markov Probabilities Implemented in Model

Source: Author's Estimates

T	T+1					Sum
	Corn	Soy	Wheat	WSG	Residual	
Corn	0.3	0.55	0.05	0.05	0.05	1
Soy	0.5	0.25	0.15	0.05	0.05	1
Wheat	0.05	0.2	0.5	0.2	0.05	1
WSG	0.05	0.05	0.15	0.7	0.05	1
Residual	0.05	0.05	0.05	0.05	0.8	1

Table 2.4: Percent Mean Absolute Error by Crop, Model Aggregation Comparison

Source: Author's Estimates

Crop	CRD	County	CRD-Cty
Corn	0.062	0.076	-0.014
Soy	0.071	0.068	0.003
Wheat	0.222	0.151	0.071
WSG	0.37	0.651	-0.281
Residual	0.047	0.052	-0.005

Table 2.5: Model Percent Mean Absolute Errors by Crop, Model Constraint Comparison  
 Source: Author's Estimates

Crop	Constrained	Unconstrained	U-C
Corn	0.235	0.062	-0.173
Soy	0.111	0.071	-0.04
Wheat	0.26	0.222	-0.038
WSG	0.994	0.37	-0.624
Residual	0.188	0.047	-0.141

Table 2.6: Regional Acreage in Million Acres by Crop Across Base and Shock Scenarios, 2001-2009 Average

Source: Author's Estimates

	WSG		Corn	Corn
	Base	Double	Upper	Lower
Corn	58.26	9.43	68.43	47.34
Soy	44.84	9.14	35.22	56.09
Wheat	20.93	8.01	17.21	27.96
WSG	30.38	69.74	23.21	42.32

### **III. A Dynamic Estimation of an Ill-Posed U.S. Feed Demand System with Distiller's Dried Grains**

#### **III.1 Introduction**

In recent years, the United States feed market has experienced compositional changes. The Energy Independence and Security Act (EISA) of 2007 includes mandates for the use of several categories of biofuel, of which some part can be met using conventional (corn starch based) and cellulosic biofuels over the next decade. After EISA passed, the amount of ethanol produced from corn has increased. This increase in corn ethanol production has also increased the availability of a co-product, distiller's dried grains (DDG). DDG are used as a feed input for various livestock types, and they have become more widely used in recent years. This increase has changed the nature of feed demands, increasing choices available to farmers for feedstocks. Because of this change, re-examining the relationships between feedstocks and explanatory data such as prices is useful, particularly to policymakers. The inclusion of DDG data would make this re-examination more pertinent to answering policy questions. This research performs such an analysis, examining the historical growth of DDG in response to changes in its own price and the prices of other feedstocks. These feedstocks may be complements or substitutes; the relationships are not initially clear.

While DDG data are available as far back as 1992, there has been a change in the market structure, with substantially larger production starting around 2002. Figure 3.1 shows U.S. DDG production from 1992-2011 as taken from U.S. Department of Agriculture, Economic Research Service (USDA/ERS). Because of this increase in production, data prior to 2002 may not be appropriate to explain the current relationships

between DDG and other feedstocks. Ergo, the data are relatively limited, and a more frequently used method of econometric analysis, such as Ordinary Least Squares (OLS), may not be appropriate. A method which is robust with limited data, or is able to handle ill-posed problems, may provide a better understanding of the relationship between DDG and traditional feedstocks. To wit, this research will assume feed use patterns follow a first-order Markov Decision Process (MDP). It is a dynamic process in which the current feed allocation is a function of the last period's allocation and a set of non-stationary transition probabilities. These probabilities are estimated based on data which affect the decision process.

Thus, the questions this research intends to answer are 1) "What relationships exist between DDG and other major animal feedstocks in the U.S.?" and 2) "Given current policy, price, and production expectations, what sort of allocation can one expect to occur with respect to DDG and other feedstock use?" To answer these questions, this research estimates feed and residual use of DDG, corn, soy meal, and an aggregation of other feedstocks using a maximum entropy estimated MDP, then projects that use forward using expected price, production, and policy data.

### **III.2 Prior Literature**

There is some literature with the estimation of feed models. Several academic and government entities produce models which include feed components. The FAPRI Baseline at the University of Missouri (Westhoff and Brown 2013) and the ERS Baseline at USDA/ERS (2013) both utilize a partial equilibrium (PE) model. A PE comprises a series of equations estimating various supply and demand functions in various sectors of the economy, varying price or trade quantities to equilibrate the model. Earlier PE models

include Womack (1976) and Adams (1994). The modeler may adjust coefficients to match known economic theory or include estimates from elsewhere in the literature. One can also implement policy shocks fairly easily by manipulating demand or supply amounts. However, while specifics may vary across specific models, PE models require parameters from some source, such as least-squares or maximum likelihood estimation, or expert advice. Estimation is difficult in the case of DDGs given the limited number of observations in the pertinent dataset, and this research offers an alternative method to estimate parameters.

Another means of constructing demand systems is an iterative linear program (LP), as is found in the implementation of POLYSYS at the University of Tennessee (Ray et al. 1998) uses iterative linear programming. In an LP, a linear production function is optimized subject to constraints such as total feedstock availability. Heady (1954) provides a discussion of the logic and advantages of using linear programs in agricultural econometrics. Tompkin (1958) uses linear programming to determine the optimal combination of production activities on a livestock farm. However, an LP may not be appropriate for this research because of the possibility of corner solutions.

GTAP at Purdue includes an implicit feed component in the oilseed/coarse grain to livestock sector interaction (Narayanan, Dimaranan, and McDougall 2012). GTAP, as a general equilibrium model, begins with a series of social accounting matrices (SAM). A SAM is a model in which different economic transactions are represented as flows between various sectors of an economy. SAMs originate with the Cambridge Growth Project (Stone and Brown 1962). While SAMs have been used for feed indirectly, they may not be appropriate for this research, given that it focuses on a single demand market.



The general equilibrium approach goes farther than social accounting, but the representation of feed demands in these models presumably rests on parameters governing price response. General equilibrium model construction likely faces the same challenge to identify key DDG-related feed elasticities using few observations that are confronted when building partial equilibrium models.

One early use of a MDP is found in Burnham (1973). The author assumes that land use changes in the Southern Mississippi Alluvial Valley can be estimated by a stationary first order Markov chain. However, he states in the footnotes that assuming stationary transition probabilities may be too restrictive for a land use change model. Burnham's conclusion is supported by Hallberg (1969), who studies frozen dairy products in Pennsylvania using a non-stationary MDP. Hallberg uses multiple regression techniques to test the hypothesis that non-stationary Markov transition probabilities have better predictive capability than stationary probabilities. MDPs are used in far more than land use, however. Again, Hallberg's study is on firm size in the dairy industry, not land use. Other studies using MDPs for econometric analysis include Adelman's analysis of the distribution of firm sizes in the iron and steel industry (1958); Paap and Van Dijk's analysis of income and consumption in the U.S. (2003); Kelley and Weiss' study of population migration based on wage differentiation (1969); Miller and Plantinga's paper analyzing land use changes in Iowa (1999); and Lubowski, Plantinga, and Stavins' use of a nested logit model to analyze national non-federal land use (2008).

One way that an estimated model can be compared to previous estimates is by looking at the differences between the price elasticities of demand. However, literature containing own- or cross-price feed demand elasticities is limited. Surry (1990) provides

an estimation of European feed demand, and those elasticities are used by GTAP (See Keeney and Hertel 2005). However, the parameters provided are the trade-off between different feedstuffs, not price elasticities, and are therefore difficult to compare to elasticities estimated by this research. Gallagher et al. (2003) provide demand elasticities for gluten feed and gluten meal in their estimation of renewable fuel standards' effect on the U.S. ethanol industry. The FAPRI Baseline (Westhoff and Brown 2013) does have elasticities, which are published in an earlier model documentation (Gerlt and Westhoff 2011). However, the representation is complicated by estimating DDG inclusion rates by type of livestock, as well as being included as part of a larger interrelated system. Westcott and Norton (2012) provide an implicit own price corn feed demand elasticity of -.20. Finally, Matthews and McConnell (2012) provide own and cross price elasticities for each feedstock studied here as part of their discussion of the US market for feed proteins, but some data of their sample period predate the rising role of DDG in feed markets. A more in-depth discussion of these elasticities accompanies the model's estimated elasticities later in the paper.

The MDP is probabilistic in nature and uses shares of total feed and residual use. Therefore, it is prudent to look at other methods of share estimation. Most of the prior literature focuses on land shares, not feedstock shares, but the estimation methods are the same. One model used to estimate land use shares and transition probabilities is the multinomial logit model (MNL) (Theil 1969). The MNL is used by McRae (1977), among others. The probabilities associated with a change in state are estimated using a logistical form, with coefficients estimated using ML. This functions similarly to a Markov chain, but disregards the prior state, estimating changes in land use solely as a

function of exogenous data. Such use can be seen in Wu and Segerson (1995); Hardie and Parks (1997); and Ahn, Plantinga, and Alig (2000), for instance. In addition, Lubowski, Plantinga, and Stavins (2003, 2008) note that the Independence from Irrelevant Alternatives (IIA) property of MNLs may preclude otherwise optimal choice behaviors. They, in addition to Lubowski (2002), use a model known as a nested logit model (NLM) to address this shortcoming of the MNL. The NLM separates decision states into subgroups or “nests” of similar qualities, differentiating them based on degree of substitutability. The nests Lubowski, Plantinga, and Stavins use include urban, non-farm (comprised of forest and range land), and farm (comprised of cropland, Conservation Reserve Program land, and pasture land). The advantage of the NLM is that it imposes IIA within nests, but not across nests, relaxing the choice restrictions. However, because the nests of a NLM are based on substitutability, it may be less efficient at explaining feed use change between feedstocks. The mixed logit model (MLM) is another method of relaxing the IIA restriction. The MLM is given by McFadden and Train (2000), and takes the choice specific variables of conditional logit and the choice-independent variables from the MNL to create a mixed model in which additional choices change the relative probabilities of the existing states.

However, OLS estimated models, MLM, and their variations still require a problem that is well posed: i.e. the observations available exceed the number of unknowns. In the case of the increasing use of DDG, the pertinent known data are very limited, and the problem may be ill-posed. Golan, Judge, and Miller (1996) discuss the use of general maximum entropy to address the issue of an ill-posed problem. From this discussion, a model that estimates a set of coefficients directly linking explanatory state

variables (both state dependent and independent to avoid IIA) with decision variables, as well as incorporating data from consumption in the prior period may be derived to provide a more robust analysis. One such model that fits those criteria is the maximum entropy estimated MDP model.

The question arises of whether or not an MDP model is appropriate for a feed model. While no literature was found using Markov chains with regard to feed, several were found using Markov models with respect to the livestock industry. Chan (1981) applies a Markov model to the Canadian cattle industry. Disney, Duffy, and Hardy (1988) use Markov chains to analyze pork farm sizes. Kristensen and Jørgensen (2000) use a Markov chain to address livestock herd management issues. Azzam and Azzam (1991) use an MDP to examine livestock with both spring and fall calving.

The MDP described below is an application to feed demand, and is motivated thusly: corn, soy meal, and DDG feed demand, each as a share of total feed demand, will be estimated as a function of own and cross prices, an index of animal numbers, and other variables included to address limitations in available data. In this representation, the feed use in the current period is affected by the allocation in the prior period. The use of shares is related to the idea that animals have a certain total feed requirement, and feedstocks can only comprise a certain percent of their diets, based on nutritional needs. Applying this to a model with short-term path dependency might include information about animal nutritional needs that is not explicitly stated. DDGs have differing rates of substitutability for existing feedstocks among various livestock types: beef cattle on feed and breeding swine may use them most readily, followed by dairy cattle and market swine, and then poultry. Without explicit variables detailing the dietary needs of each

livestock type, lagged shares may indirectly include information about feedstock inclusion rates for the animals represented by the animal units index. The successes and failures of this functional form in explaining changes in feed use demand are detailed in the following sections.

### III.3 Markovian Feed Use Decision Process

This research examines the dynamics of feed use changes by explaining the decisions of feed mixing companies in the area of study. Following Ahn, Plantinga, and Alig (2000), who focus on land allocation rather than livestock feed, each company in region  $i$  ( $i = 1, \dots, I$ ) is assumed to purchase a sequence of feedstocks that maximizes the present discounted value of expected net returns

$$(3.1) \quad \max_{\delta_{j,k,t}} \sum_{t=0}^{\infty} \tau^t \cdot E[NR(\delta_{j,k,t}, \mathbf{X}_{h,i,t}) + \varepsilon_t(\delta_{j,k,t})]$$

where  $\tau$  is a constant discount factor,  $\delta_{j,k,t}$  represents a decision to allocate funds dedicated to the purchase of feedstock (discrete state variable)  $j$  in time  $t-1$  to feedstock  $k$  in time  $t$ ,  $\mathbf{X}_{h,i,t}$  represents a  $T \times H$  matrix of observable data, and  $\varepsilon_t$  represents unobserved variables. To state it another way, a mixing company manager purchases a series of inputs to produce feed mixes using funds in the budget. This allocation of funds for the purchase of inputs changes over time, and the decision is impacted by the observable data, such as input prices. Beyond budget limits faced by the mixing company, this maximization is unconstrained. Because of the relative difficulty involved with obtaining individual company decisions, a model which utilizes aggregated data is desirable. Therefore, aggregated feed and residual disappearance is used to represent the sum of all purchase decisions, in a manner that is intended to be analogous to previous applications

to the land use problem (Golan, Judge, and Miller 1996). However, there exist idiosyncratic or agent-specific components of a purchase decision, which are unobservable to the econometrician. These components are represented by the unobservable state variable  $\varepsilon$ . Over time, a feed mixing company's purchase decisions are assumed to follow a first order MDP. The amount of feed inputs purchased from a particular stock is a function of the amount purchased in the prior period and a non-stationary transition probability, noted  $\pi_{i,j,k,t}$ . That is, there is a non-observable  $J \times K$  (where  $J=K$ ) matrix of probabilities for each state and time period, transitioning from feedstock  $j$  to feedstock  $k$  from period  $t-1$  to period  $t$ . Because of the cyclical nature of the livestock industry (Stockton and Tassell 2007), one can surmise that there is a similar path-dependent element to the feed industry supporting it. Furthermore, DDGs may substitute for corn and soy meal at differing rates for each livestock type. Hoffman and Baker (2011) provide a meta-study discussion of these differences, noting beef cattle on feed have the highest potential inclusion rates, with a range of 20-40% of dry matter intake, compared to poultry's relatively lower 10-15% potential inclusion rates for DDGs. However, it is difficult to incorporate this data explicitly in the model. However, because feed purchasers already have this information, the differentiation in inclusion rates may be represented in prior purchase decisions. This provides the motivation for using a model in which the current state is a function of the prior state.

To illustrate, Table 3.1 shows a set of probabilities for feed use transition from 2003-2004 as estimated from the model explained below. If a portion of a feed mixing company's total quantity of feed purchases was dedicated to purchasing corn in 2003, it has approximately a 45% chance of staying in corn in 2004, a 7.5% chance of being used

to buy soy meal, a less than 1% chance of being used for DDG, and a 46% chance of being used for some other feedstock. These numbers are only present to provide a visual example of the nature of Markovian transition probabilities; as results, they will be discussed later. Markovian transition probabilities row-sum to one and are hypothesized to be affected by explanatory variables including feedstock prices. Through this interaction, the variables are explanatory of changes in feedstock purchase patterns.

The model is derived using maximum entropy (ME), following applications to the land share allocation problem as seen in Golan and Vogel (2000); Golan, Judge, and Miller (1996); and Miller and Plantinga (1999). The ME method for estimating the model of Markov transition probabilities is set forth by Jaynes (1957). The objective of the ME method is to select the probabilities that use the least information (fewest assumptions) to estimate the probabilities while still satisfying the constraints. This, in turn, assumes the mixing company has the greatest amount of choice possible. Shannon's (1948) entropy measure is used to measure the amount of information needed to estimate the coefficients here for feed demand allocation, as it has been in the context of the land allocation problem (Miller and Plantinga 1999, Golan and Vogel 2000). The primal objective function determining the optimal transition probabilities of the feed demand allocation problem is assumed to be

$$(3.2) \quad \begin{aligned} \max_{\pi_{i,j,k,t}} S_i &= \sum_{t=1}^T \sum_{j=1}^J \sum_{k=1}^K \pi_{i,j,k,t} \cdot \ln(\pi_{i,j,k,t}) \\ s.t. \quad \sum_{t=1}^T \mathbf{X}'_{i,t} \left( y_{i,k,t} - \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t} \right) &= \bar{0} \end{aligned} ,$$

where  $\bar{0}$  is an  $H$  size vector of zeroes and  $y_{i,k,t}$  is the share of feed purchased from stock  $k$  in area  $i$  during time  $t$ . When applied to the estimating equations (the set of constraints), the solution to the problem is assumed to take the form

$$(3.3) \quad \hat{\pi}_{i,j,k,t} = \frac{q_{i,j,k,t} \cdot \exp\left(y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \hat{\lambda}_{h,i,k}\right)}{\sum_{k=1}^K q_{i,j,k,t} \cdot \exp\left(y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \hat{\lambda}_{h,i,k}\right)},$$

where  $\hat{\lambda}_{h,i,k}$  is the optimal Lagrangian multiplier associated with explanatory variable  $h$  and stock  $k$ , and  $q_{i,j,k,t}$  is a conditional probability, which may be adjusted by the analyst to represent information from before the purchase decision that may bias the mixing company's decision. Here, priors are uniform, and the mixing company is assumed to have no prior bias to its purchase decision. By focusing on the dual of the primal problem, an unconstrained equation to determine the optimal multipliers takes the form of

$$(3.4) \quad \max M_i(\bar{\lambda}) = \sum_{t=1}^T \sum_{h=1}^H \sum_{k=1}^K y_{i,k,t} x_{h,i,t} \lambda_{h,i,k} - \sum_{t=1}^T \sum_{j=1}^K \ln \left[ \sum_{k=1}^K q_{i,j,k,t} \exp \left( y_{i,j,t-1} \cdot \sum_{h=1}^H x_{h,i,t} \cdot \lambda_{h,i,k} \right) \right].$$

One  $k$  category is kept as a residual, with its multipliers assumed to be zero. Because of the additive nature of the feedstock categories, the residual solution is implicit when all other categories are estimated. After the multipliers are estimated, they are used with the other variables to determine the transition probabilities in (3.3). The transition probabilities are then applied to the prior year's feedstock purchases to determine current purchases:

$$(3.5) \quad \hat{y}_{i,k,t} = \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t}.$$



With the feed shares estimated, elasticities can be calculated to show the effect of a one percent change in the explanatory variables on the transition probabilities. The impacts are calculated in the same method used in land allocation problems (Miller and Plantinga 1999), namely

$$(3.6) \quad \omega_{h,i,j,k,t} = x_{h,i,t} \cdot y_{i,j,t-1} \left[ \hat{\lambda}_{h,i,k} - \sum_{k=1}^K \hat{\lambda}_{h,i,k} \cdot \pi_{i,j,k,t} \right]$$

Using this equation, one can derive the feedstock elasticities, measuring the change in the feed allocation from a one percent change in the explanatory variables.

$$(3.7) \quad \kappa_{h,i,k,t} = \hat{y}_{i,k,t}^{-1} \sum_{j=1}^J y_{i,j,t-1} \cdot \pi_{i,j,k,t} \cdot \omega_{h,i,j,k,t}$$

Using the appendix relating to the land allocation problem in Miller and Plantinga (1999) as a guide, the covariance matrix of the coefficients ( $\hat{\lambda}_{h,i,k}$ ) conditional on the explanatory variables can be estimated by

$$(3.8) \quad \left( (\mathbf{X} \otimes \mathbf{I}_{(K-1)^2})' \boldsymbol{\Sigma}_i (\mathbf{X} \otimes \mathbf{I}_{(K-1)^2}) \right)^{-1}$$

such that  $\mathbf{I}$  is the identity matrix, and  $\boldsymbol{\Sigma}_i$  is a  $T(K-1)^2 \times T(K-1)^2$  matrix where the diagonal elements are  $\pi_{i,j,k,t} (1 - \pi_{i,j,k,t})$  and the off diagonal elements are  $\pi_{i,j,k,t} (-\pi_{i,j,k,t})$ .

With the MDP model defined, the focus shifts to defining the data that will be used in the model.

#### III.4 Data: Feed and Residual Use

This research is concerned with the effects of the increase in DDG use for feed. However, precise feed use data are difficult to obtain. USDA/ERS (Capehart, Allen, and Bond 2013; Ash 2013) is one of the primary sources for feed disappearance data, but their estimation of grain and oilseed meals allocated to feed also include a residual

component. Which is to say that they have data on supply, including domestic production, imports, and beginning stocks. They also have data on most demands, including exports, domestic food/industrial use, and ending stocks. To make supply and demand balance, the remaining demand is assumed to be taken by feed and residual use. Feed use, therefore, is not directly estimated, but rather calculated as an accounting balance. This calculation and the problems it presents are discussed in detail in Westcott and Norton (2012). To address the residual component of feed data, they estimate using feed and residual as the left-hand-side variable, but include variables which, while one might not include in a demand system, are correlated with the causes of residual variation, and will ostensibly capture residual effects. The two variables they use are corn production, with the assumption that larger corn crops increase the likelihood of shrinkage and estimation errors, and the percent of the corn crop rated mature by August 31 in the next year, with the assumption that a larger early harvest increases the likelihood of new crop usage, resulting in estimation errors of ending stocks. As such, both of these variables are included in the **X** matrix for this research. While this may not provide precise feed use estimates, it should mitigate the residual effects in the data. Because one category of shares is known as “residual”, “feed and residual use” will hereafter be referred to simply as “feed use”, with the residual component of the data taken as a given for clarity.

The next step is to discuss the actual feed data gathered. Corn and soy meal, and hay are the major feedstocks in the U.S., with DDG and various other grains and meals following. National level feed use was gathered for each of these from 2002-2012, converted into million tons, and summed to calculate total feed and residual use. The

MDP requires a lag year, and so the sample period runs ten marketing years from 2003-2012. The feedstocks, means, and standard deviations are given in Table 3.2. The data for hay, wheat, sorghum, canola meal, oats, barley, cottonseed meal, sunflower meal, linseed meal, and fish meal are aggregated into the residual category. Hay disappearance is added to the residual category for two reasons. One, the author's previous experience with this method has shown that a large residual category is desirable: when returns to all other uses increase, the residual category is from where the increases in those uses come. As such, a residual category that is large enough to handle large shifts in other categories allows the model to work better. The annual use of the other stocks only summed to seventeen million tons on average, considered very small against annual corn feed use of almost 150 million tons on average. The addition of hay to the residual category increased it by almost 145 million tons on average, allowing for a much greater transition "buffer". Two, this research is less concerned with DDG effects on hay usage as with effects on corn and soy meal usage.

### **III.5 Data: Explanatory Variables**

There are several factors which affect the choices for feed use. A traditional demand system includes own and cross price effects, as well as other demand drivers. For this research, the prices of each non-residual category are included as explanatory variables. Grain consuming animal units (GCAU) are included as a demand driver. Following the discussion above from Westcott and Norton (2012), corn production and percent of corn maturity by September 1 are included. Table 3.2 lists the variables included along with their means and standard deviations. An intercept term is also included. All prices listed are in real terms. All data come from ERS and NASS.

Current prices, as opposed to lagged prices, are used in the model. This makes sense because purchase decisions are made using current price information instead of expected prices, as a farmer might use for the end of the growing season. However, at this level of aggregation, it presents a problem of endogeneity. That is, changes in aggregate consumption patterns will likely have an effect on prices. While this problem is known, it is difficult to find suitable instruments for feedstock prices at this level of aggregation: it is suspected, but unconfirmed, that the United States' status as a net agricultural exporter in corn, soy meal, and DDGs may cause a similar correlation in the prices of other countries. In such a scenario, the endogeneity bias of the estimated coefficients would not be removed, merely modified. Therefore, current prices are included as independent variables, recognizing the possible bias present in the estimators.

### III.6 In Sample Dynamic Validation

The model is run with the listed variables over the sample period, starting with known data in the initial lag year, and estimating iteratively over the course of the sample period. Figures 3.2, 3.3, and 3.4 show the historical shares compared against the model estimates. The lines represent historical shares, and the points represent the model estimates. The model estimates track very closely with historical estimates, showing a very high goodness-of-fit. As far as a numerical measure of goodness-of-fit, typical  $R^2$ 's cannot be calculated for this method. However, a goodness of fit measure can be calculated by taking the errors in the estimates, normalized by the share amount, and averaging them over time:

$$(3.9) \quad MAE_k = \frac{1}{TI} \sum_{t=1}^T \sum_{i=1}^I \frac{abs(y_{i,t,k} - \hat{y}_{i,t,k})}{y_{i,t,k}}.$$

MAE is the percent mean absolute error: the difference of the estimated share allocated to a particular feedstock and the actual share over time for each feedstock, normalized by the actual share allocated to that feedstock, averaged over time. As it measures deviation from historical data, a lower number indicates a better fit. Because  $y$  represents a share, as opposed to actual acreage, this estimate is a unitless percentage. This does not have the same statistical application as a normal  $R^2$ , but it does provide an idea of how the model performs. Table 3.4 shows the measures of fit, and they support the graphical validation of the model: the average errors are very low. The model performs worse in estimating DDG, but it still performs well overall.

The last statistical measure is the errors associated with the feedstocks and explanatory variables. The results of (3.8) are presented in Table 3.5. Two variables proved statistically significant: corn price and DDG price, both with regards to DDG feed use, with corn price more significant than DDG price. While the other explanatory variables did not prove to be statistically significant at the 10% level of significance, they are economically significant, as their inclusion has been explained prior.

### **III.7 Model Elasticities**

The model estimates feedstock elasticities from (3.7) above. These results are presented in Table 3.6. The first thing to note is that the own price elasticities are all negative, so the model seems to follow economic theory at first glance. The negative elasticity on soybean meal price to corn feed suggests that corn and soy meal are complimentary, and this is echoed in the soy meal to corn price elasticity. The magnitudes are different, but the lack of statistical significance in each of the coefficients may indicate a large error in the estimates. Symmetry is not imposed in these estimates.

The positive cross price elasticity on DDG feed use to corn indicates that DDG will compete with corn for market share, and this is echoed in the positive elasticity on DDG prices to corn use. DDG's relationship with soy meal is less clear. There is a positive effect in soy meal use to DDG price, suggesting competition, but the opposing effect, DDG use to soy meal price, is negative, and very small in absolute value. There is literature (Hoffman and Baker 2011, e.g.) suggesting that DDGs substitute for soy meal as well as corn, in differing ratios based on the need of the animals in question. It is the author's belief, based on this questionable result, that the nature or amounts of DDGs substituting for soy meal changed over the course of the sample period. This is a possible explanation, but not certain. The corn and soy meal uses are price inelastic, both own and cross, but DDG are much more elastic.

It is important to note here that the average elasticities shown here for DDGs do not tell the whole story. The DDG own price elasticities are dynamic over time. Figure 3.5 shows the DDG price elasticities to each feedstock over the in-sample period. This figure is important for three reasons: one, it shows that corn and soy meal responses are fairly flat over time. This suggests that feedlot owners have a fairly well established idea of how they will respond to changes in price. The slow increase in the elasticities of these two feed inputs might be caused by the growing volumes of DDG that could have given feed mills more options, and led to greater price responsiveness. Two, it shows that DDG price responses are very dynamic, changing in relatively large amounts from year to year. This suggests that DDGs are still being "felt out" in the sample period. This supports the prior claim of a changing market structure. It further shows that a model which relies on limited data, such as this one, may be robust with regard to DDG estimation. Three, the

trend in DDG own-price responsiveness is increasing in absolute terms throughout the sample period. Because the share of total feed use is increasing during that time, this is somewhat counterintuitive. One would expect responsiveness to decrease as quantity demanded increases. The exact cause of this is difficult to pinpoint. However, the increasing share might have a mitigating effect on the rising elasticity of demand, and that the actual price responsiveness could be higher than the elasticity alone states.

GCAU elasticities follow expectations with corn and soy meal, showing a positive elasticity. DDG response to GCAU does not follow theory as expected: one would expect a positive sign, since DDGs are often fed to animals which are heavily weighted in the calculation of GCAU, such as beef cattle (Berger 2008). However, beef cattle production decreased over the sample period, decreasing the main source of DDG feed use represented in GCAU, potentially causing this seemingly odd result.

As mentioned prior, there is some literature with which one can compare these elasticities. However, care must be taken when so doing. The majority of prior literature does not provide directly comparable elasticities: either different measures of demand are used, or different elasticities (quantity comparisons vs. response to price, e.g.) are given. Moreover, this research estimates based on shares of total feed demand, instead of gross feed in terms of bushels or metric tons. However, even barring direct comparison, one can gain some sense of how this model functions relative to previous works. The most direct comparison one can make is with Westcott and Norton (2012), who provide an implicit corn price to corn feed demand elasticity of  $-0.20$ . This is relatively similar to the model's estimate of  $-0.61$ : both show corn feed demand to be relatively price inelastic, although this model is more responsive than Westcott and Norton's. GTAP uses

substitution elasticities – the trade-off between feed inputs in quantity terms – from Surry (1990). These are presented in Table 3.7 for easy comparison. It should be noted that Surry models European feed demand, not US. This has particular implications with the substitution for corn, wheat, and barley. As is seen in Table 3.2, corn is used substantially more than wheat or barley for feed in the US. In addition, Surry does not include soy meal, but he does include a “high-protein” category, which may be used to compare. That said, the own substitution elasticities may correlate to an own price elasticity, because it considers at what rate a firm owner would substitute a feedstuff for itself, which would rely chiefly on own price. One can see that corn is much more elastic with Surry than this model. Soy meal is likewise more elastic with Surry than this model. DDGs are not included in Surry and cannot be compared. Gallagher et al. (2003) provide elasticities for gluten meal (high protein) and gluten feed, both by products of wet corn milling. Gluten feed can be compared to DDG, although DDG are produced in dry mill plants as opposed to wet mill plants (Hoffman and Baker 2010). The own price elasticity from Gallagher is very strong relative to the one estimated by this model: -31.5 vs. -2.37. This may be explained, however, by the substantially smaller amount of DDGs consumed in 2000, which is Gallagher’s baseline year, versus current amounts, reflected in the average presented here.

The FAPRI Baseline (Westhoff and Brown 2013) published elasticities may not be directly compared, because they are included as part of a larger endogenous system, with multiple related equations. That is, a single equation elasticity will not give the actual impact of effects in price changes on feedstock use. The author obtained DDG model impact multipliers, estimated by exogenizing prices and livestock numbers and



shocking the system with a 10% change to the prices in question. The elasticity-analogous system impact multipliers from the 2013 FAPRI-MU DDG feed model are shown in Table 3.8. FAPRI corn price impacts on DDG feed demand are similar to this model's in that they are both relatively strongly positive, however the FAPRI estimate is smaller in magnitude. The reverse occurs with DDG own price response: DDG feed use is less sensitive to own price changes in this model than in the FAPRI model. The elasticity for DDG use to soy meal price is substantially different with differing signs as well as magnitude. However, issues arising from soy-DDG interaction have already been discussed.

Elasticities are also found in Matthews and McConnell (2012), and are presented in Table 3.9.<sup>7</sup> This is the most comprehensive comparable set of elasticities found, although it should be noted that the variables are defined slightly differently in that paper, and cover a much broader period (1992-2010). Grain represents feed grains in general, corn, barley, sorghum, and oats inclusive. Oilseed meal represents all protein meals fed, where as many of those are kept in the residual category here. Co-products represents DDGs as well as corn gluten feed. The elasticities are likewise negative for own price, but all their cross price elasticities are positive, implying strict substitution. There are a couple differences that stand out. First, their DDG response to corn price is much, much smaller than the one estimated by this model. This may be attributed to the inclusion of much older data, when DDG use was much lower than it is currently. The relatively flat and low DDG use in the early years may cause reactions to changes in corn price to be understated. Second, their own price elasticities are also relatively smaller for corn and

---

<sup>7</sup> Matthews and McConnell also present elasticities for urea and meat/bone meal, which are not of concern here. They are left out of the table in question.

DDGs. Again, this decreased responsiveness may be explained by the inclusion of older data.

One can see that the model elasticities fall fairly close to those in the established literature, DDG-soy interaction excepted, and, more importantly, the own price elasticities agree with regard to sign across all other studies. Moreover, one might point out that the elasticities estimated in the model are estimated without ad-hoc adjustment. While one may make the argument for constraints on this model, the estimation procedure is explored without constraints here, as befits the competition/substitution question is difficult to answer with regard to soy meal.

### **III.8 Projection Data**

With the model validated and following economic theory, the focus turns to projecting DDG use over the next ten years. The FAPRI Baseline provides estimates which may be used for projection purposes. There was a small issue with GCAU calculation. The FAPRI Baseline animal units measurements are scaled differently than those of ERS. To scale the FAPRI projections to match the model, the ERS GCAU and HPAU historical numbers were estimated as a function of the FAPRI historical estimates and projected forward based on FAPRI projected estimates.<sup>8</sup> Corn production was taken directly from the baseline. The percent corn early harvest in  $t+1$  appeared to be a stationary function, and was projected based on the mean of the sample period.

---

<sup>8</sup> There are other ways of approaching this issue, such as using FAPRI estimates throughout, or basing the forecast of the ERS series on other projections. The decision to use USDA data in historic estimation reflects the desire to rely on publicly available data for the historical period to make results more readily reproduced. For the projections, however, the FAPRI-MU baseline projections were more readily available and up to date at the time of writing.

### **III.9 Projection Shocks and Results**

Using the projection data described above, the model is run out of sample from 2013-2022, using known initial lag shares from 2012, providing a ten year projection of feed share use. Along with this estimate, two additional scenarios are run, with projection DDG prices at 125% and 75% of base projection amounts. This is included to both provide a broader range of possible results and to include the possibility of policy shocks, such as a modification in the EISA mandates, which would change ethanol and DDG production, changing DDG prices. The results of the projections are shown in Figures 3.6, 3.7, 3.8, and 3.9. The first thing to note is that the DDG estimates continue to follow economic theory, with higher prices resulting in lower feed demand. Corn shares follow similarly, with higher cross prices resulting in higher own use.

DDG's relationship with soy meal is less clear, however. The overall effects are very small, with soy meal use changing less than 2% of total feed demand over the projection period, but the scale has been reduced in Figure 3.8 to better show the impact. Soy meal begins as expected, assuming competition, with higher DDG prices resulting in more soy meal usage. In 2014, this relationship changes, with DDG and soy meal acting as complements for the rest of the projection. This is a questionable result, and, due to the relatively small magnitude of the effects and the lack of significant variables, it is difficult to pin down an exact cause or explanation for this result. In an attempt to explain this, the model was backed up an additional three years, and dynamically simulated starting from 2010, to see if the change from in- to out-of-sample had an appreciable impact. This is shown in Figure 3.8a. This does not appear to shed much light on the situation, but there is one particular thing which stands out: the simultaneous cross

at 2014 does not occur, broken apart to 2013 and 2012. This suggests that there is a path dependency issue present. Because the MDP is a single year iterative model, this may have a large impact on results. Further study was necessary to clarify exactly what happened. Given the dynamic nature of DDG own price elasticities shown in Figure 3.5, DDG price to soy meal use elasticities were examined in the projection period, presented in Figure 3.9, along with DDG prices in real terms. The start year of 2013 has a visible impact on the out-of-sample estimation because of the spike in DDG prices in 2012. This causes a strong increase in the elasticities, which, combined with the small shares of soy meal in the feed market, causes the first year results to be questionable. To put it another way, it might be the case that the early out-of-sample results are a lagged outcome of the price spike in the final in-sample year. However, since the price decreases over time, the initial strong response simply corrects itself over the course of the projection period, and the model reaches a stable state.

The more interesting story comes from the direction and changes in DDG use. DDG use decreases over the whole of the projection period, with about 5% of total feed use being taken out of DDGs, decreasing in the baseline and high price scenarios. This decrease goes chiefly to corn, which gains about 6% across each scenario.

### **III.10 Conclusion**

This research intended to examine the relationships between DDG and traditional feedstocks. DDG were found to respond as expected to economic theory using an unconstrained, historically validated statistical model, with negative own- and mixed cross-price elasticities. This suggests that DDG is a substitute to corn used for feed, but the interaction with soy meal is unclear.. This research also sought to examine what the

DDG market might look like in the next ten years. To answer this, the existing relationships in the model were projected forward using price and production estimates from the FAPRI Baseline. Under these conditions, DDG feed use is expected to decrease slightly over the next ten years, with an increase in corn feed. Soy meal projections were flat in the base case.

DDG and corn feed demand responses in the event of a higher or lower DDG price followed expectations. For example, a higher DDG price would lead to lower DDG feed demand in the projections and more corn feed use. The impacts of DDG price on soy meal were very small, but directions were initially difficult to interpret. It was hypothesized that the relatively small share of soy meal and dynamic DDG price to soy meal use elasticities contributed to this issue. However, soy meal was estimated to act as a complement to DDGs over the projected long term. The method used here shows some promise for unconstrained estimation of feed demands over a short sample period in order to take account of the recent increase in DDG volumes. The key challenges identified here include a theoretical one of aligning land allocation applications to the present feed demand problem more reliably. Endogeneity of right hand side variables introduces risks in estimation that are not easily addressed without reliable instruments. Of course, data are always imperfect, and the “feed and residual” category controls implemented here might not correct this particular problem. Nevertheless, estimated elasticities are often similar to those identified by other researchers and the method explored here merits further investigation.

### III.11 References

- Adams, G. 1994. "Impact Multipliers of the U.S. Crops Sector: A Focus on the Effects of Commodity Interaction." PhD dissertation, University of Missouri.
- Adelman, I.G. 1958. "A Stochastic Analysis of the Size Distribution of Firms." *Journal of the American Statistical Association* 53: pp. 893- 904.
- Ahn, S., A.J. Plantinga, and R.J. Alig. 2000. "Predicting Future Forestland Area: A Comparison of Econometric Approaches." *Forest Science* 46(3): pp. 363-376.
- Ash, M. 2013. "Oil Crops Outlook." Washington DC: U.S. Department of Agriculture, Econ. Res. Service, OCS-13I, September.
- Azzam, S.M., and A.M. Azzam. 1991. "A Markovian Decision Model for Beef Cattle Replacement That Considers Spring and Fall Calving." *Journal of Animal Science* 69: pp. 2329-2341.
- Berger, L.L. 2008. "Distiller's Grain in Beef Cattle Diets." Presentation, Univ. of Illinois, Urbana-Champaign. March 31.
- Burnham, B.O. 1973. "Markov Intertemporal Land Use Simulation Model." *Southern Journal of Agricultural Economics* 5(1): pp. 253-258.
- Capehart, T., E. Allen, and J.K. Bond. 2013. "Feed Outlook." Washington DC: U.S. Department of Agriculture, Econ. Res. Service, FDS-13I, September.
- Chan, M.W. L. 1981. "A Markovian Approach to the Study of the Canadian Cattle Industry." *Review of Economics and Statistics* 63(1): pp. 107-116.
- Disney, W.T., P.A. Duffy, and W.E. Hardy, Jr. 1988. "A Markov Chain Analysis of Pork Farm Size Distributions in the South." *Southern Journal of Agricultural Economics* 20(2): pp. 57-64.
- Gallagher, P.W., H. Shapouri, J. Price, G. Schamel, and H. Brubaker. 2003. "Some Long-Run Effects of Growing Markets and Renewable Fuel Standards on Additives Markets and the US Ethanol Industry." *Journal of Policy Modeling* 25: pp. 585-608.
- Golan, A., G. Judge, and D. Miller. 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York NY: John Wiley and Sons.
- Golan, A., and S. Vogel. 2000. "Estimation of Non-Stationary Social Accounting Matrix Coefficients with Supply Side Information." *Economic Systems Research* 12(4): pp. 447-471.
- Hallberg, M. 1969. "Projecting the Size Distribution of Agricultural Firms: An Application of a Markov Process with Non-Stationary Transition Probabilities." *American Journal of Agricultural Economics* 51: pp. 289-302.

- Hardie, I.W., and P.J. Parks. 1997. "Land Use with Heterogeneous Land Quality: An Application of an Area Base Model." *American Journal of Agricultural Economics* 79(2): pp. 299-310.
- Heady, E. O. 1954. "Simplified Presentation and Logical Aspects of Linear Programming Technique." *Journal of Farm Economics* 36(5): pp. 1035- 1048.
- Hoffman, L.A., and A. Baker. 2010. "Market Issues and Prospects for U.S. Distillers' Grains: Supply, Use, and Price Relationships." Washington DC: U.S. Department of Agriculture, Econ. Res. Service, FDS-10k-01, December.
- . 2011. "Estimating the Substitution of Distillers' Grains for Corn and Soybean Meal in the U.S. Feed Complex." Washington DC: U.S. Department of Agriculture, Econ. Res. Service, FDS-11-I-01, October.
- Jaynes, E.T. 1957. "Information Theory and Statistical Mechanics." *Physical Review* 106(4): pp. 620-630.
- Keeney, R. and T.W. Hertel. 2005. "GTAP-AGR: A Framework for Assessing the Implications of Multilateral Changes in Agricultural Policies." GTAP Technical Paper No. 24, Center for Global Trade Analysis, Purdue University, West Lafayette, IN. August.
- Kelley, A.C. and L.W. Weiss. 1969. "Markov Processes and Economic Analysis: The Case of Migration." *Econometrica* 37(2): pp. 280-297.
- Kristensen, A.R. and R. Jørgensen. 2000. "Multi-Level Hierarchic Markov Processes as a Framework for Herd Management Support." *Annals of Operation Research* 94: pp. 69-89.
- Lubowski, R.N. 2002. "Determinants of Land-Use Transitions in the United States: Econometric Analysis of Changes Among the Major Land-Use Categories." PhD Dissertation, Harvard University.
- Lubowski, R.N., A.J. Plantinga, and R.N. Stavins. 2003. "Determinants of Land Use Change in the United States 1982-1997." Discussion Paper 03-47, Washington DC: Resources for the Future, August.
- . 2008. "What Drives Land-Use Change in the United States? A National Analysis of Landowner Decisions." *Land Economics* 84 (4): pp. 529-550.
- Matthews, K.H., Jr., and M.J. McConnell. 2012. "The Market for U.S. Livestock Feed Proteins." *Applied Economic Perspectives and Policy* 34(4): pp. 555-569.
- McFadden, D., and K. Train. 2000. "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* 15: pp. 447-470.
- McRae, E. 1977. "Estimation of Time-Varying Markov Processes with Aggregate Data." *Econometrica* 45: pp. 183-98.

- Miller, D.J. and A.J. Plantinga. 1999. "Modeling Land Use Decisions with Aggregate Data." *American Journal of Agricultural Economics* 81(1): pp. 180-194.
- Narayanan, B., B.V. Dimaranan, and R.A. McDougall. 2012. "Chapter 2: Guide to the GTAP Data Base." GTAP 8 Data Base Documentation, Center for Global Trade Analysis, Purdue University, West Lafayette, IN.
- Paap, R. and H.K. van Dijk. 2003. "Bayes Estimates of Markov Trends in Possibly Cointegrated Series: An Application to U.S. Consumption and Income." *Journal of Business & Economic Statistics* 21: pp. 547-563.
- Public Law (PL) 110–140. 2007. Energy Independence and Security Act of 2007. [frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi? dbname=110\\_cong\\_public\\_laws&docid=f:publ140.110.pdf](http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ140.110.pdf). March 10, 2008.
- Ray, D.E., D.G. De La Torre Ugarte, M.R. Dicks, and K.H. Tiller. 1998. "The POLYSYS Modeling Framework: A Documentation." Staff Paper, Agricultural Policy Analysis Center, Dept. of Agricultural Economics and Rural Sociology, University of Tennessee.
- Shannon, C.E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27: pp. 379-423, 623-656.
- Stockton, M. C., and L. W. Van Tassell. 2007. "The Cattle Price Cycle: An Exploration in Simulation." Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management. Chicago IL.
- Surry, Y. 1990. "Econometric Modeling of the European Compound Feed Sector: An Application to France," *Journal of Agricultural Economics*, 41: pp. 404-421.
- Stone, R. and A. Brown. 1962. "A Computable Model for Economic Growth." Cambridge UK: Cambridge Growth Project.
- Theil, H. 1969. "A Multinomial Extension of the Linear Logit Model." *International Economic Review* 10: pp. 251-9.
- Tompkin, J.R. 1958. "Response of the Farm Production Unit as a Whole to Prices Response of the Farm Production Unit as a Whole to Prices." *Journal of Farm Economics* 40(5): pp. 1115-1128
- US Department of Agriculture, National Agricultural Statistics Service (USDA/NASS). 2012. "Crop Production 2011 Summary." Available at [usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047](http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1047).
- Westcott, P.C. and J.D. Norton. 2012. "Implications of an Early Corn Crop Harvest for Feed and Residual Use Estimates." Washington DC: U.S. Department of Agriculture, Econ. Res. Service, FDS-12f-01, July.



- Westhoff, P., and S. Brown. 2013. "US Baseline Briefing Book: Projections for agricultural and biofuel markets." MU-FAPRI Report #01-13, University of Missouri.
- Womack, A. 1976. "The U.S. Demand for Corn, Sorghum, Oats, and Barley: An Econometric Analysis." Economic Report 7615, Department of Agriculture and Applied Economics, University of Minnesota, St. Paul.
- Wu, J., and K. Segerson. 1995. "The Impact of Policies and Land Characteristics on Potential Groundwater Pollution in Wisconsin." *American Journal of Agricultural Economics* 77(4): pp. 1033-1047.

### III.12 Appendix: Figures and Tables

Figure 3.1: DDG Production 1992 – 2011

Source: ERS Feedgrains Yearbook

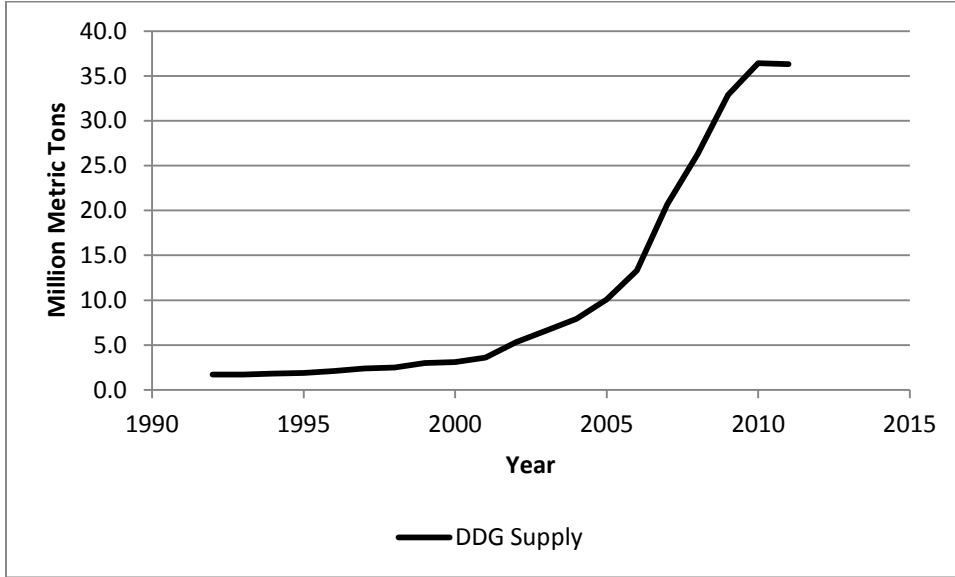


Figure 3.2: Corn Feed and Residual Use as a Percent of Total Feed and Residual Use, Historical Amounts versus Dynamic Model Estimates, 2003-2012

Source: ERS and Author's Estimates

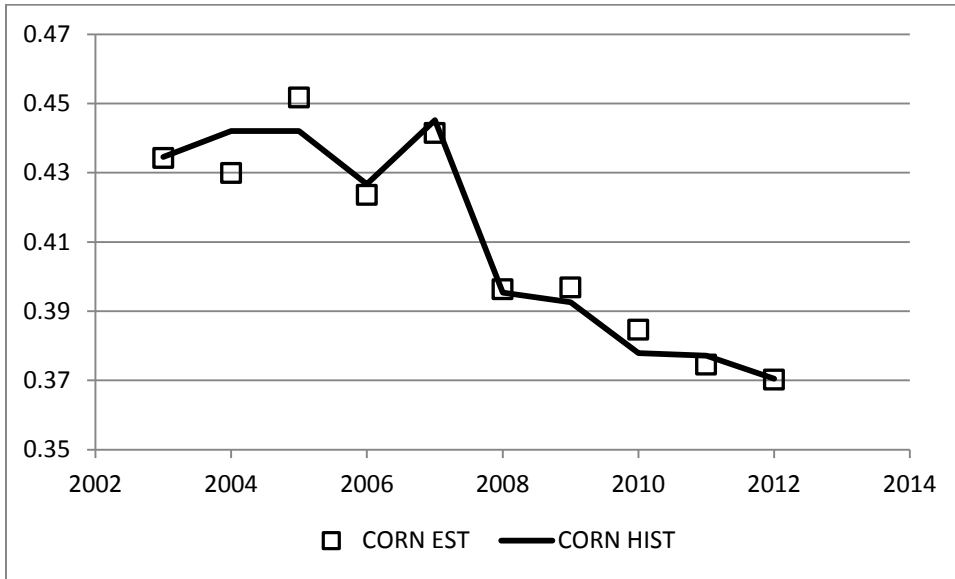


Figure 3.3: Soy Meal Feed and Residual Use as a Percent of Total Feed and Residual Use, Historical Amounts versus Dynamic Model Estimates, 2003-2012  
 Source: ERS and Author's Estimates

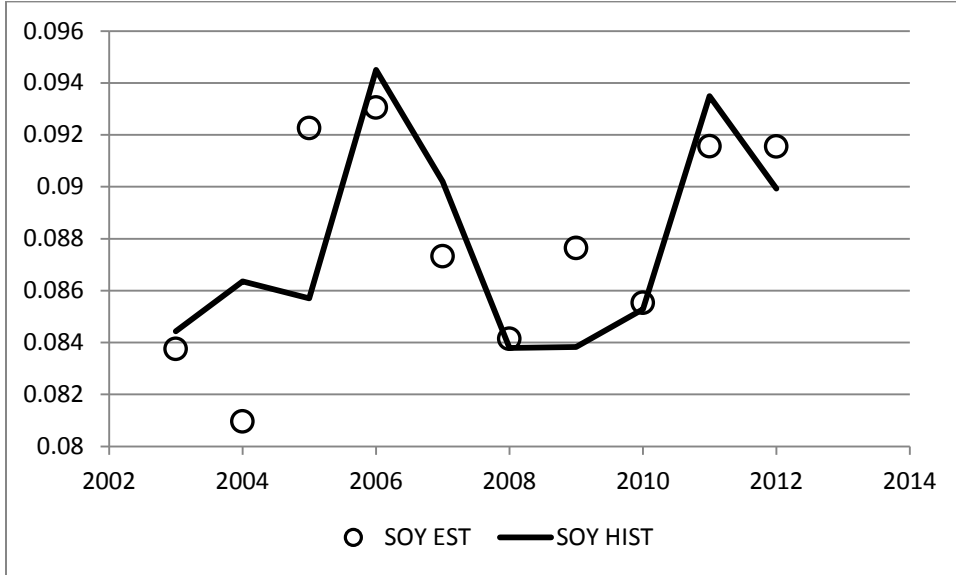


Figure 3.4: DDG Feed and Residual Use as a Percent of Total Feed and Residual Use, Historical Amounts versus Dynamic Model Estimates, 2003-2012  
 Source: ERS and Author's Estimates

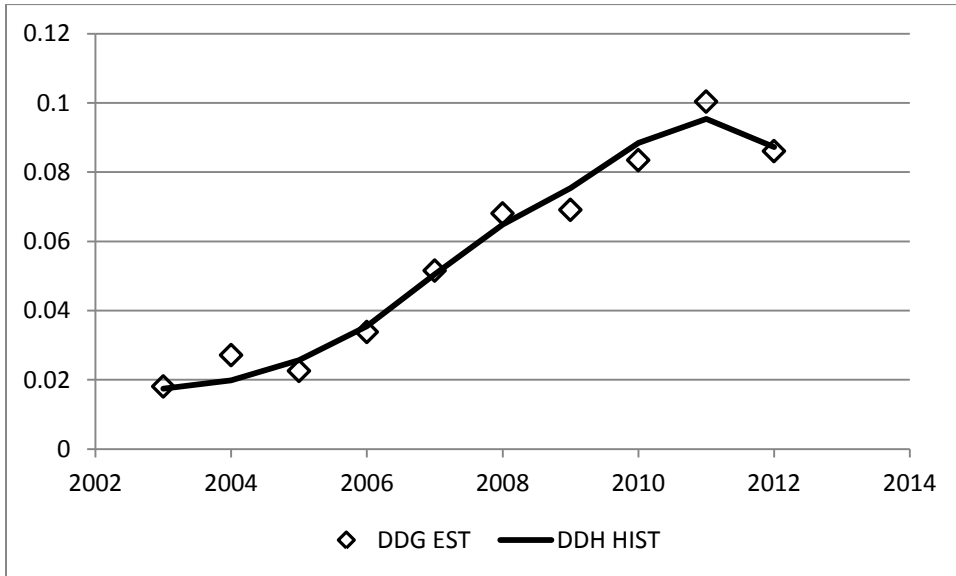


Figure 3.5: DDG Price to Feed and Residual Use (as a Percent of Total Feed and Residual Use) Elasticities by Feedstock, 2003-2012  
 Source: Author's Estimates

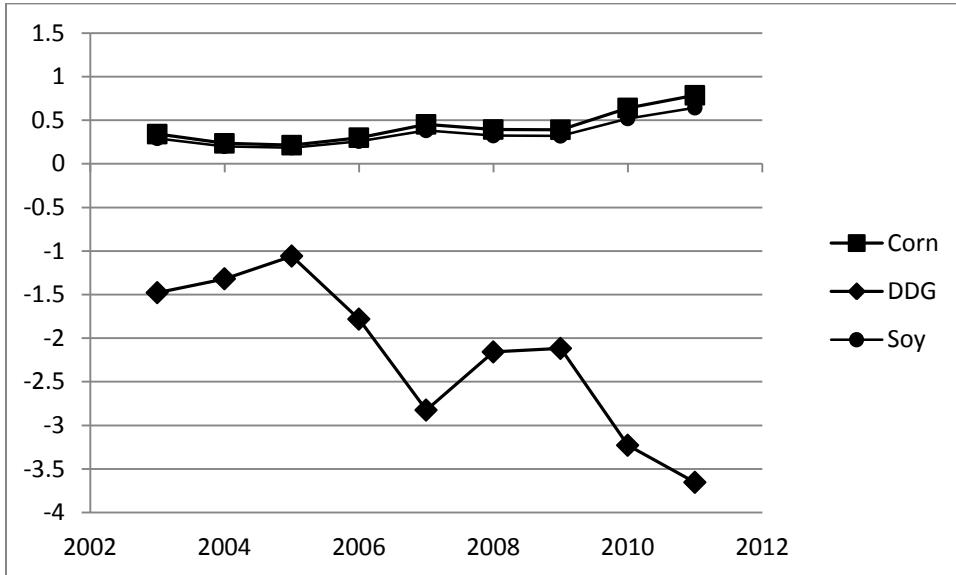


Figure 3.6: DDG Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2013-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks

Source: Author's Estimates based on FAPRI Baseline price projections.

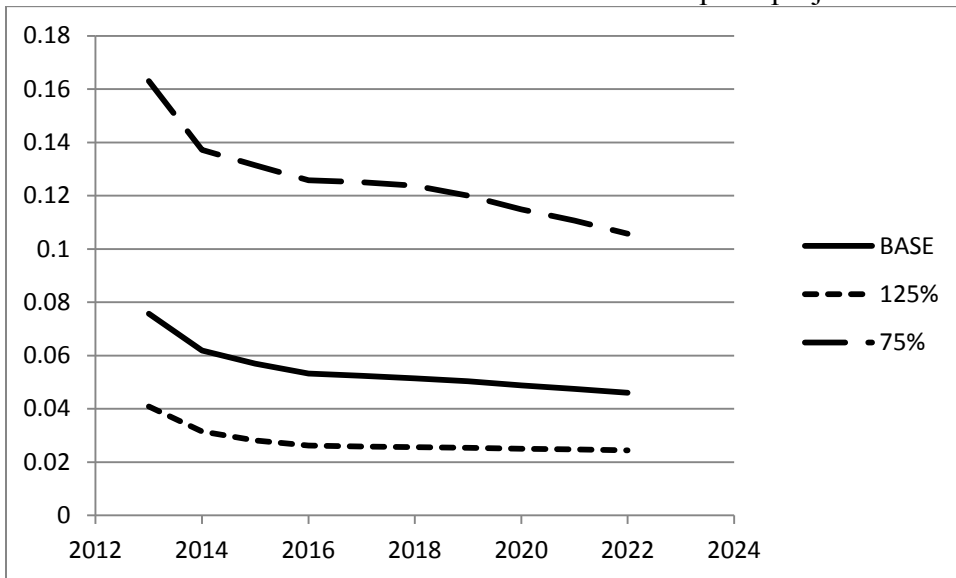


Figure 3.7: Corn Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2013-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks

Source: Author's Estimates based on FAPRI Baseline price projections.

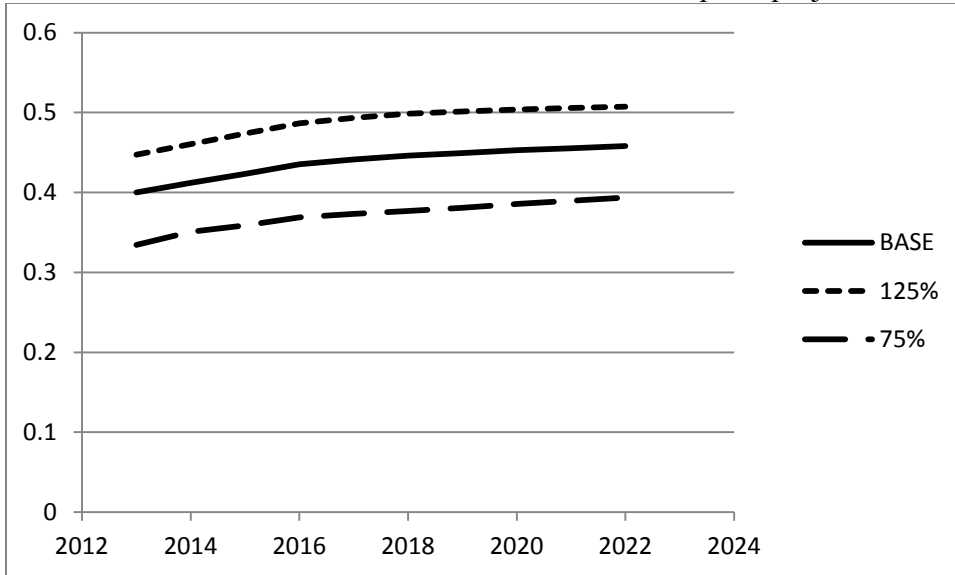


Figure 3.8: Soy Meal Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2013-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks

Source: Author's Estimates based on FAPRI Baseline price projections.

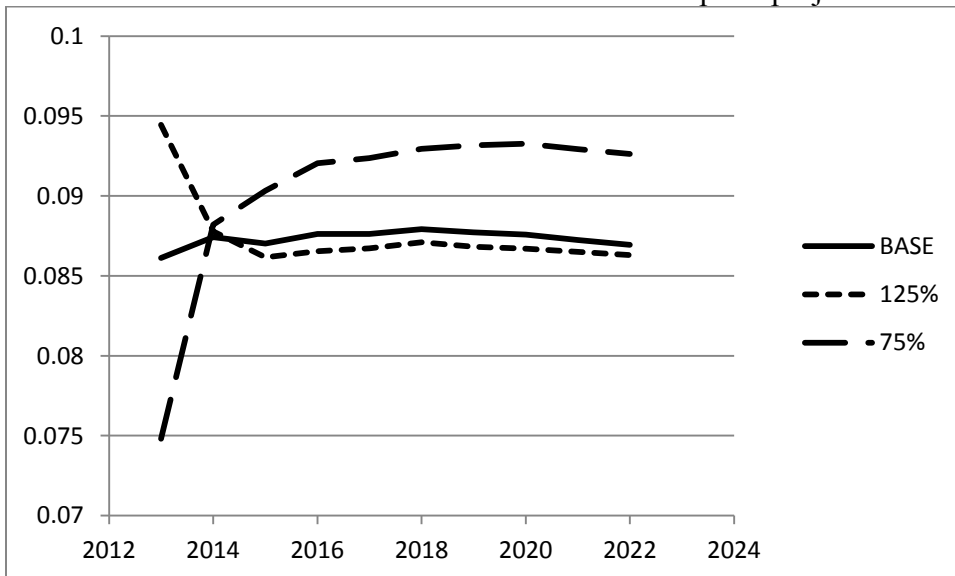


Figure 3.8a: Soy Meal Feed and Residual Use as a Percent of Total Feed and Residual Use Projections, 2010-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks

Source: Author's Estimates based on FAPRI Baseline price projections.

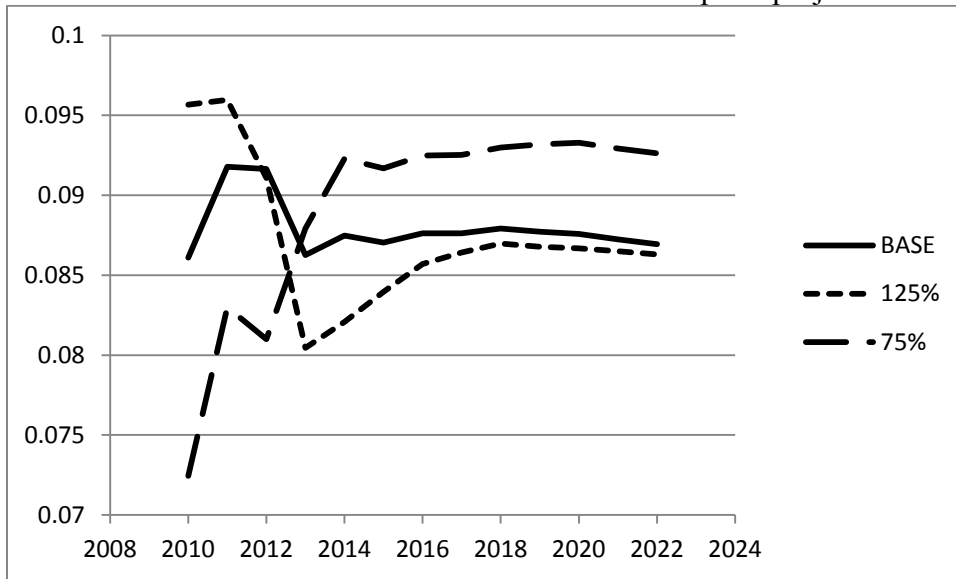


Figure 3.9: Soy Meal Feed and Residual Use (as a Percent of Total Feed and Residual Use) to DDG Price Elasticities, 2010-2022, Baseline, High DDG Price (125%), and Low DDG Price (75%) Shocks

Source: Author's Estimates

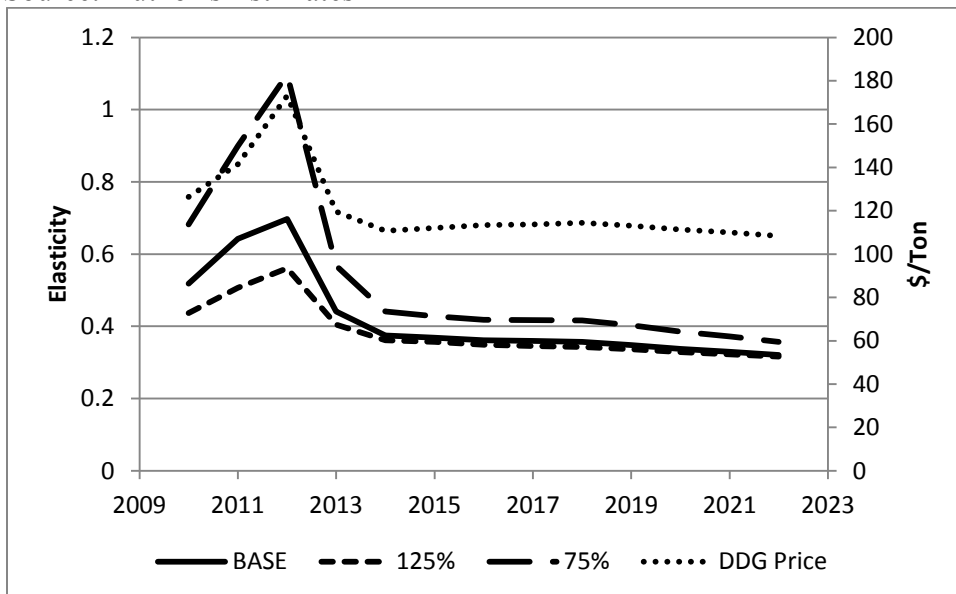


Table 3.1: 2003/04 – 2004/05 MDP Transition Probabilities

Source: Author's Estimates

2003\2004	Corn	DDG	Soy Meal	Residual	Sum
Corn	0.452	0.009	0.075	0.464	1.000
DDG	0.264	0.225	0.246	0.265	1.000
Soy Meal	0.315	0.146	0.222	0.317	1.000
Residual	0.458	0.006	0.064	0.471	1.000

Table 3.2: Feedstock Sample Mean and Standard Deviation in Million Tons

Source: ERS

Feedstock	Mean	Std. Dev
Corn	149.73	17.75
Hay	145.22	9.83
Soy Meal	31.89	1.56
DDG	19.96	10.10
Wheat	5.28	2.92
Sorghum	4.08	1.34
Canola Meal	2.63	0.72
Oats	1.87	0.31
Barley	1.37	0.50
Cottonseed Meal	1.06	0.16
Sunflower Meal	0.30	0.08
Linseed Meal	0.21	0.04
Fish Meal	0.21	0.03

Table 3.3: Explanatory Variables, Units, Means, and Standard Deviations

Source: ERS

Variable	Units	Mean	Std. Dev
Grain Consuming Animal Units	Mil Units	91.77	1.90
Corn Crop Maturity, Week of Sept 1, $t+1$	Percent	20.00	10.67
Corn Production	Mil Bu	11482.83	1303.64
Corn Grain Price	\$/Bu	2.57	0.51
Soy Meal Price	\$/Ton	198.57	29.37
DDG Price	\$/Ton	94.25	19.00

Table 3.4: Percent Mean Absolute Errors by Crop, 2003-2012

Source: Author's Estimates

Crop	MAE
Corn	0.011
DDG	0.092
Soy Meal	0.026

Table 3.5: T-Statistics of Model Coefficient Estimates by Feedstock  
Source: Author's Estimates

	Corn	DDGs	Soy Meal
Corn Price	-0.520	3.780*	-0.443
DDG Price	0.341	-2.364**	0.240
Soy Meal Price	-0.045	-0.095	-0.115
GCAU	0.304	-0.018	0.519
Corn Maturity	-0.001	0.253	0.256
Corn Production	0.107	0.621	-0.065

\* Statistically Significant at  $\alpha=0.05$

\*\* Statistically Significant at  $\alpha=0.10$

Table 3.6: Model Estimate Elasticities, 2003-2012 Average  
Source: Author's Estimates

	Corn	Soy Meal	DDG
Corn Price	-0.613	-0.517	3.491
Soy Price	-0.027	-0.094	-0.065
DDG Price	0.451	0.370	-2.367
GCAU	1.708	2.922	-1.258

Table 3.7: Feedstuff Substitution Elasticities  
Source: Surry (1990)

Feedstuff	Wheat	Corn	Barley	High-Protein	Brans	Share
Wheat	-11.93	2.94	9.22	-0.46	2.47	0.17
Corn	2.94	-1.76	-0.15	0.15	0.92	0.35
Barley	9.24	-0.15	-22.21	0.81	3.36	0.09
High-Protein	-0.46	0.15	0.81	-0.33	0.40	0.28
Brans	2.47	0.93	3.36	0.41	-11.55	0.10

Table 3.8: FAPRI Baseline Elasticity-Analogous System Impact Multipliers for DDG Feed Use

Source: Westhoff and Brown (2013) and Wyatt Thompson, FAPRI-MU

	DDG Feed Use
Corn Price	2.17
Soy Meal Price	1.34
DDG Price	-3.57



Table 3.9: Feed Use Elasticities by Feedstock, 1992-2010  
 Source: Matthews and McConnell (2012)

Feed Quantities	Prices		
	Corn Price	Soy Meal Price	DDG Price
Grains	-0.139	0.125	0.002
Oilseed Meals	0.296	-0.354	0.030
Co-products	0.024	0.168	-0.210

## VITA

Dustin J. Donahue was born in Chattanooga, TN in 1982. He graduated with honors from Ooltewah High School in May 2000, and continued on to the University of Tennessee, Chattanooga. He received a B.S. in Economics in December 2004 and was accepted in Omicron Delta Epsilon. His work there focused on tax theory in cooperation with E. Bruce Hutchinson and J.R. Clark. He continued his education at the University of Tennessee, Knoxville, graduating with a M.S. in Agricultural Economics. While there, he was accepted to Gamma Sigma Delta and worked with APAC, focusing on biofuel production econometrics. He was awarded a USDA National Needs Fellowship and completed his PhD in Agricultural and Applied Economics at the University of Missouri in December 2013. While there, he worked with the Food and Agricultural Policy Research Institute on areas including biofuel policy and land-use allocation problems. His current plans are to work in academia, focusing on teaching college-level economics.